Regulated complex assembly protects cells from aberrant *Sleeping Beauty* transposition events

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)

im Fach Biologie
eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät I

der Humboldt-Universität zu Berlin

von
Diana Pryputniewicz-Drobińska

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I
Prof. Dr. Lutz-Helmut Schön

Gutachter/innen:

1. Prof. Dr. Thomas Börner
2. Prof. Dr. Achim Leutz
3. Prof. Dr. Gerald Schumann

Tag der mündlichen Prüfung: 08.03.2010

ZUSAMMENFASSUNG

Transposons sind genetische Elemente, die fähig sind, sich innerhalb des Genoms zu bewegen. Obwohl Transposons lange Zeit als parasitäre Elemente oder gar molekularer Müll galten, erfahren sie seit einiger Zeit Anerkennung als Werkzeuge für unterschiedliche genetische Anwendungen. Sleeping Beauty (SB) gehört zur Tc1/mariner-Superfamilie von Transposons. SB wurde aus molekularen Fossilien rekonstruiert [Ivics, 1997], um u.a. einen sicheren und effizienten Vektor für die Gentherapie zu schaffen. Zu diesem Zweck ist es notwendig, den Mechanismus der SB-Transposition und deren Regulation, die Aktivitäten des Proteins und den Einfluss von Wirtsfaktoren genau zu verstehen. Aus diesem Grund habe ich in meiner Arbeit die einzelnen Schritte des Transpositionsprozesses und die Bildung des sogenannten paired-end complex (PEC) – eine Voraussetzung für die folgenden katalytischen Reaktionen – untersucht. Zusätzlich habe ich versucht, einen in vitro Transpositionstest für SB zu etablieren. Ein solcher Test könnte der schnellen Untersuchung von Mechanismen dienen, die die Regulation der Transposition beeinflussen.
SB gehört zur IR/DR-Gruppe der Tc1/mariner-Superfamilie von Transposons. Im Gegensatz zu mariner-like-Elementen mit kurzen inverted repeats (IR) ist die IR/DR-Struktur von SB durch lange IRs mit insgesamt vier Bindestellen für die Transposase gekennzeichnet. Ich habe die Fähigkeit dieser beiden Transposon-Systeme zum Ausschneiden eines Transposonendes ohne die Beteiligung des anderen Endes im PEC getestet. Solche unpräzise Transposition kann zu genomic rearrangements führen. Meine Ergebnisse zeigen, dass SB zwar imstande ist, ein einzelnes Transposonende auszuschneiden, dies geschieht jedoch weit weniger effizient als bei mariner-like-Elementen. Die Unterdrückung unpräziser Transpositionsereignisse ist ein Ergebnis der besseren Regulation von SBs Transposition, die durch die IR/DR-Struktur bedingt ist. Die Komplexität der IRs in Kombination mit der zweiteiligen (PAIRED) DNA-Bindedomäne von SB kann als Mittel einer raffinierten Regulation des Transpositionsprozesses angesehen werden, welche das Genom vor anormalen Transpositionsereignissen schützt. Die Ergebnisse meiner Arbeit legen ein Modell nahe, in dem die Bildung des PEC während der Transposition von SB ein höchst genau regulierter Prozess ist, der durch die DNA-Protein- und Protein-Protein-Bindeaffinitäten der beiden DNA-Binde-Subdomänen geleitet wird.

ABSTRACT

Transposons are pieces of DNA able to move within the genomes. These mobile elements previously seen as “junk DNA” gained much interests recently as possible tools for various genetic applications. Sleeping Beauty is a verterbrate Tc1/mariner transposon reconstructed from molecular fossils [Ivics, 1997] to create a safe and efficient vector for gene therapy. For that purpose it is important to deeply understand the mechanism and regulation of the SB transposition, the activities of the transposase and influence of host factors on the process. Therefore, in this project I studied the single steps of the transposition reaction and formation of the paired-end complex (PEC) which is a prerequisite for the subsequent catalytic steps. Additionally, I tried to establish an in vitro transposition assay for Sleeping Beauty that would serve an easy assay for testing the system and probe mechanisms affecting the regulation of transposition activity.
Sleeping Beauty belongs to the IR/DR subfamily of the Tc1/mariner-like transposons. In contrast to mariner-like elements with short inverted repeats (IR), the IR/DR structure of SB is characterized by long IRs with four binding sites for the transposase. I compared the ability of the two systems to perform cleavage of the single transposon end without including the second end in the PEC. Such imprecise transposition can lead to genome rearrangements. My results show that SB is capable of single-end cleavage; however, to much lower extent than the mariner-like element. Lower number of imprecise transposition events is a result of better regulation of the SB transposition that might be imposed by the IR/DR stucture. The complexity of the inverted repeats together with the bipartite (PAIRED) DNA-binding domain of SB might offer means for more sophisticated regulation of the transposition process, thereby protecting the genome from aberrant transposition events. I propose that complex formation in SB transposition is a strictly regulated ordered assembly process, guided by DNA-protein and protein-protein interaction interfaces of the DNA-binding subdomains. Obtained results allowed me to draw a model how the paired-end complex is formed.

ABBREVIATIONS

ATP	adenosine triphosphate
bp	base pair
C-/N-terminus	carboxy-/amino-terminus
cpm	counts per minute
DNA	desoxyribonucleic acid
DR	direct repeat
E.coli	*Escherichia coli*
HMGB1	High-mobility group box 1
HTH	helix-turn-helix
IHF	integration host factor
IR	inverted repeat
IS	insertion sequence
kb	kilobase
kD	kilodalton
LINE	long-terminal interspersed elements
LTR	Long Terminal Repeat
M	molar
MBP	maltose-binding protein
min	minute
neo	neomycin
NLS	nuclear localization signal
OD	optical density
ORF	open reading frame
PCR	polymerase chain reaction
PEC	paired-end complex
RNA	ribonucleic acid
RNAi	RNA interference
SB	Sleeping Beauty
SINE	short interspersed elements
TE	transposable element
wt	wild type
X-SCID	X-linked severe combined immunodeficiency

1 INTRODUCTION

1.1 Transposable elements

Transposable elements are pieces of DNA capable of moving from one place to another within the genome. First transposons were discovered in maize (*Zea mays*) by Barbara McClintock in the 1940s [McClintock, 1950]. Studying the mechanisms of the mosaic colour patterns of maize seeds, she discovered that two genetic loci responsible for the mosaicism are able to change their positions on the chromosomes. She named the loci *Dissociator* (*Ds*) for causing the breakage of the chromosome 9 and *Activator* (*Ac)* whose presence was required for chromosome dissociation.

Nowadays, we know that transposons are widespread in nature, present in bacteria, invertebrates, vertebrates and plants. They constitute about 45% of the human genome [Lander, 2001]. Even though, for long time considered as useless "junk DNA", studies showed that transposons have evolutionary impact and their presence has been found in ~4% of human protein-coding regions [Nekrutenko, 2001], [Volff, 2006]. Additionally, transposable elements gained much attention as possible tools for practical applications, e.g. gene therapy.

Transposons relocate in a recombination process called transposition. Transposition is the recombination reaction that mediates movement of DNA segments between many nonhomologous sites. Because the transposable elements are mobile, they have the capacity to change host gene information, for example, by inserting into a gene and disrupting its coding sequence or by inserting adjacent to a gene and influencing its expression. Transposons play a very important role in evolution being a source of many genetic variations.

Different types of the mobile elements have been identified and classified based on their structure and mechanism of transposition. There are two main classes of transposable elements depending on the intermediate used in the transposition reaction. Class I gathers elements utilizing RNA as the intermediate and thus called retroelements; class II elements use DNA as the transposition intermediate (Abb. 1). DNA transposons are common in bacteria, invertebrates and plants, whereas in vertebrates retrotransposons are more abundant [Henikoff, 1992]. Retroelements contain a reverse transcriptase gene and transpose duplicatively in a process called retroposition. The DNA element is transcribed to RNA, then reverse-transcribed to DNA, and integrated elsewhere in the genome. Retroelements are divided into two major groups depending on the presence or absence of Long Terminal Repeats (LTR). The non-LTR elements including short interspersed elements (SINE) and long-terminal interspersed elements (LINE) make up 33.9% of the human genome (reviewed in [Bannert, 2004]). The LTR class includes retrotransposons, endogenous retroviruses and group II introns.

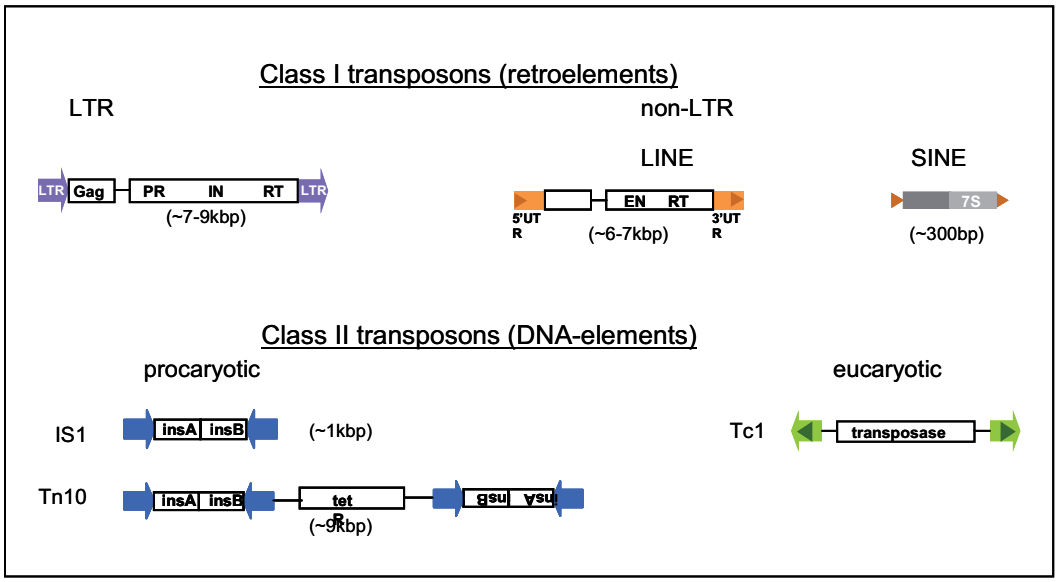

Abb. 1: Classification of transposable elements. Transposable elements are grouped into two classes depending on the intermediate used in the transposition reaction. LTR-retrotransposons have Long Terminal Repeats (LTR) flanking genes coding for gag (protein required for the assembly of the RNA transcript into cytoplasmic particles), protease (PR), integrase (IN) and reverse transcriptase (RT); LINEs (long-terminal interspersed elements) encode endonuclase (EN) and RT; SINEs (short interspersed elements) are derived from 7SL RNA and have no coding capacity; DNA-transposons consist of inverted repeats flanking gene encoding for the transposase; bacterial composite transposons (here Tn10) carry antibiotic resistance gene (tetR) flanked by full-length insertion sequences (IS).

Hundreds of DNA-transposons have been identified in bacteria. Several kinds of transposable elements can be distinguished, such as simple insertion sequences, composite transposons and bacteriophages. Majority of insertion sequences (IS) are composed of recombinationally active 10–40 bp short DNA sequences that define the ends of the element and flank an open reading frame encoding the transposase enzyme which recognizes and processes these ends. Based on similarities in genetic organization and in transposase structure, ISs have been classified into 19 families. Members of the *IS630* family show significant similarities with eukaryotic Tc1/*mariner*-like elements. Tn7 is an example of another type of bacterial transposons, that similarly to ISs has short inverted repeats, but in contrast to ISs it codes for a number of proteins, five transposition genes and two antibiotic resistance genes [Rogers, 1986].
The third group consists of composite transposons, such as Tn5 and Tn10, which DNA region coding for antibiotic resistances is flanked by IS elements at both ends in opposite orientation [Berg, 1982], [Foster, 1981]. The transposase is encoded by the right IS and is able to mobilize the ISs alone as well as the complete element [Kleckner, 1996]. Mu is a bacteriophage that infects *E.coli* using DNA-based transposition to integrate its genome into the genome of the host cell. It can then use transposition to initiate its viral DNA replication. Mu is a very complex

transposable element requiring for the process two Mu proteins, Mu A and Mu B, two host proteins, HU and IHF,[Krause, 1986], [Lavoie, 1993] and the presence of transpositional enhancer [Surette, 1989], [Chaconas, 1996]. Important feature of the Tn5 and Tn10 transposons and the Mu bacteriophage is the *trans* activity of the catalytic domain of the transposase. It means that the transposase bound to one transposon end process the cleavage of the second transposon end [Naumann, 2000]. In such case, formation of the paired-end complex bringing two transposon arms together is strictly required for catalytic steps to take place. This trait might be also shared by other prokaryotic elements.
The largest group of eukaryotic DNA-transposons is composed of members related to the *mariner* and the Tc1 families.

*1.2 Tc1/*mariner *family*

Members of the Tc1/*marine*r transposon family are found in eukaryotes ranging from insects and nematodes to fish and humans (reviewed in [Plasterk, 1999]. Tc1/*mariner* elements are about 1.3 to 2.4 kb in length and consist of a single gene encoding the transposase enzyme flanked by terminal inverted repeats (IRs). The IRs of different elements vary in length, but all of them contain sites recognized and bound by the transposases. The transposase is the protein catalyzing the remobilization of the transposable element. The protein has an N-terminal DNA-binding domain specifically recognizing the transposon DNA [Vos, 1994], [Colloms, 1994], [Wang, 1999], [Izsvak, 2002] and a C-terminal catalytic domain with a conserved DDE or DDD motif [Doak, 1994]. Transposase activity is sufficient to mediate full transposition reaction *in vitro* [Vos, 1996], [Lampe, 1996]. Nevertheless, this evidence does not exclude the involvement of host proteins *in vivo*. The sequence and length of the flanking inverted repeats are not conserved within the family, with the exception of the most terminal nucleotides (5'-CAGT) [Radice, 1994]. All Tc1/*mariner* elements insert exclusively into a TA dinucleotide leading to duplication of the TA sequences [Plasterk, 1996].
The Tc1/*mariner* family can be divided into two subfamilies. The founding element of the Tc1 family was initially detected in the *C. elegans* genome and was called Tc1 for transposon caenorhabditis number 1 [Emmons, 1983]. The first *mariner* subfamily member was identified in *Drosophila mauritiana* [Jacobson, 1985]. Since then, many elements were found in various organisms, but only few are naturally active. Most Tc1/*mariner* transposons accumulated over the time inactivating mutations, frame shifts or small deletions resulting in nonfunctional elements. In vertebrates, not a single active element has been identified. However, several active elements have been reconstructed based on sequences of multiple dead copies. The first molecular reconstruction brought to life a fish element, called *Sleeping Beauty* [Ivics, 1997]. *Sleeping Beauty* was followed by *Frog Prince* from frog species *Rana pipiens* [Miskey, 2003] and *Himar1* from the horn fly *Haematobia irritans* [Lampe, 1996].

1.3 Sleeping Beauty

Isolation of an active element from vertebrates would open a way to develop transposon-based vectors for germline transformation and insertional mutagenesis in this group of animals. As the attempts to find an active element failed, a reconstruction from molecular fossils has been undertaken. A member of the Tc1/*mariner* family was chosen, as the family seemed not to be as demanding for species-specific factors.
Sleeping Beauty transposon system was reconstructed based on phylogenetic data collected from Tc1-like elements in different teleost fish species and comparative analysis of functional transposase domains [Ivics, 1997]. The *SB* transposon system consists of two components: a synthetic gene encoding a transposase enzyme, and a nonautonomous salmonid-type element carrying the inverted repeats. This system is very similar to a transposon active approximately 10 million years ago [Ivics, 1996] that was able to invade a great number of teleost genomes.

1.3.1 The *SB* transposon system (Abb. 2)

1.3.1.1 Transposase

The *SB* transposase has an N-terminal bipartite DNA-binding domain [Vos, 1994] similar to the paired domain originally described in important developmental regulatory proteins in eukaryotes, such as the Pax family of transcription factors [Czerny, 1993], [Gehring, 1994], [Jun, 1996]. The DNA-binding domain consists of two helix-turn-helix motifs (HTH) called PAI and RED (PAI+RED = PAIRED domain). The PAI subdomain is responsible for the specific binding to the DNA binding sites [Izsvak, 2002]. The function of the second HTH (RED subdomain) has not been defined. As in all PAIRED-like domains, a GRPR-like motif (GRRR in SB) is located between the two HTH motifs [Gehring, 1994], [Izsvak, 2002]. It mediates interactions with DNA in the Hin invertase of *Salmonella* [Feng, 1994] and in the RAG1 recombinase of V(D)J recombination [Spanopoulou, 1996]. The RED subdomain overlaps partially with the nuclear localization signal (NLS). The NLS is flanked by a putative phosphorylation target sites of casein kinase II [Ivics, 1996]. Phosphorylation of these sites may play a role in the regulation of the transposition.
Catalytic domain of the *SB* transposase responsible for the DNA cleavage and joining reactions is located in the C-terminus of the protein. It contains a DDE motif consisting of two aspartate (DD) and one glutamate (E) amino-acid residues. The DDE motif is conserved among transposases, integrases and recombinases and is essential for catalytic activities. Closely related *mariner* transposases have a DDD triad and substitution of the third aspartatic residue with the glutamatic residue (D→E) results in inactivation of the protein [Lohe, 1997]. Addition-

ally, there is a glycine-rich region within the catalytic domain of the *SB* transposase. However, its function remains unknown.

1.3.1.2 Transposon DNA

Tc1-like elements, including *SB*, but not the Tc1 itself, have special organization of inverted repeats, termed IR/DR structure [Plasterk, 1999]. The inverted repeats (IR) are approximately 230 bp long and each contains two binding sites for the transposase. The binding sites are very similar in sequence and heading in the same direction, thus called direct repeats (DR). However, the DRs are not identical, showing 3 nucleotides differences in sequence and 2 nucleotides differences in length. The binding sites placed directly at the transposon ends are called outer DRs and the binding sites placed approximately 165 bp distantly from the cleavage sites are called inner DRs (Abb. 2B). All four binding sites within the IR/DR structure are required for SB transposition [Izsvák, 2000]. In contrast, for another Tc1-like element, Tc3, the internal binding sites play no major role in the transposition reaction, since deleting it does not affect the transposition frequency [Fischer, 1999]. Additionaly, the inverted repeat upstream of the *SB* transposase gene contains a transpositional enhancer-like sequence, termed HDR [Izsvak, 2002].

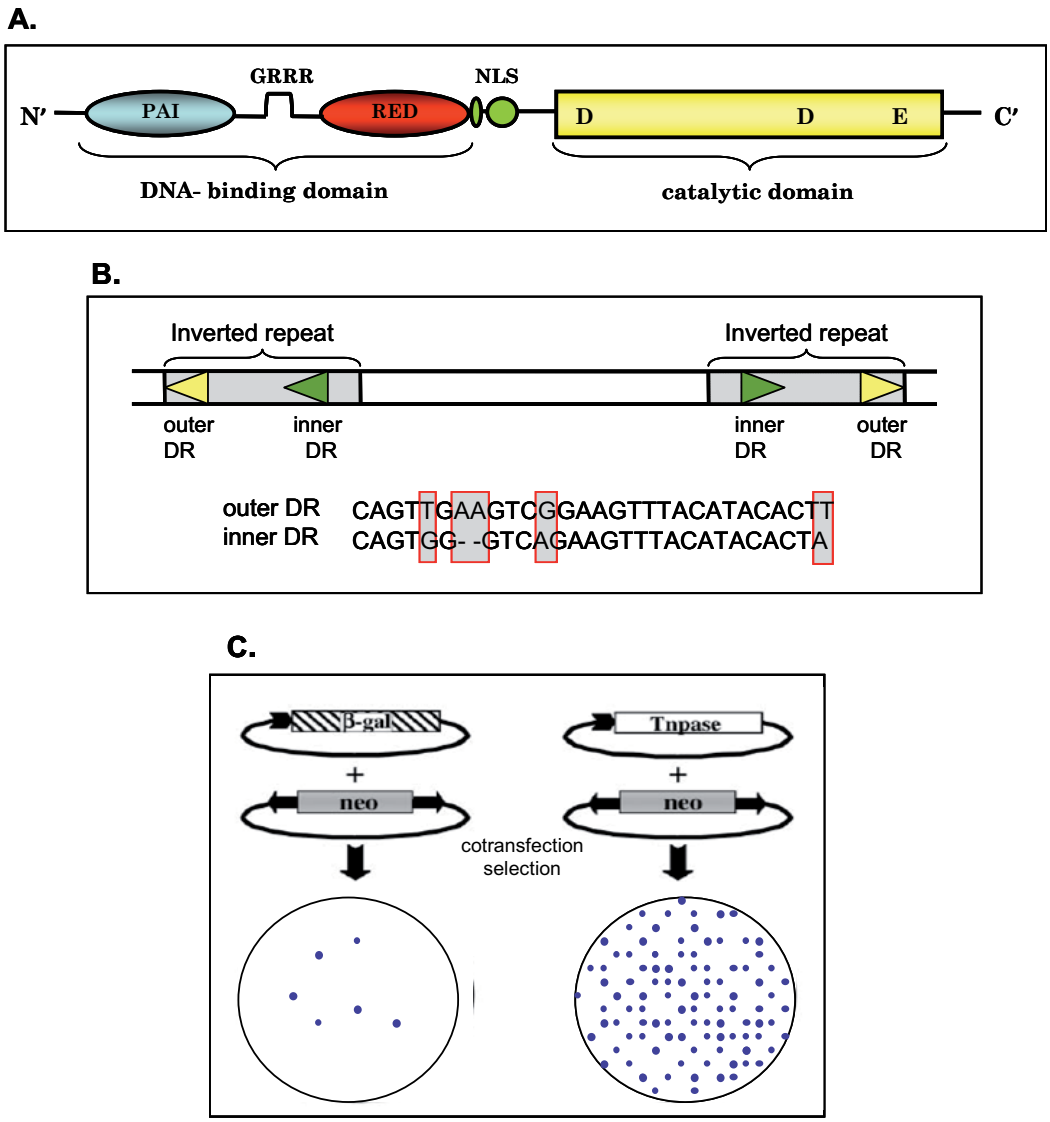

Abb. 2: The *Sleeping Beauty* transposon system. A. structure of the *SB* transposase consisting of bipartite DNA-binding domain (PAI and RED), nuclear localization signal (NLS) and C-terminal catalytic domain; **B.** structure of the *SB* transposon DNA consisting of two inverted repeats (IR). Each IR has two binding sites (DR) for the transposase (shown as yellow and green triangles); under the structure comparison of the sequences of inner and outer DR with differences highlighted in the grey boxes; **C.** *in vivo* transposition assay: cells are cotransfected with plasmid carrying the inverted repeats flanking a selection marker together with a transposase expressing plasmid or an unrelated gene as a negative control. The ratio between the numbers obtained in the presence versus the absence of transposase is the readout of the assay [**Ivics, 1997**].

An *in vivo* transposition assay was developed to check the ability of the two-component *SB* system to undergo transposition [Ivics, 1997]. Plasmid with the reconstructed transposase gene and plasmid carrying the inverted repeats flanking a selection marker (neomycin-resistance) gene were cotransfected. When the

transposition takes place and the transposon is reinserted into the genome, the cells gain the resistance and grow under neomycin selection. The formed colonies are stained and counted. The ratio between the numbers obtained in the presence versus the absence of transposase is the readout of the assay, and is the measure of the efficiency of the transposition (Abb. 2C). The transposition assay was performed in cells of different vertebrate classes: seven fish species, mouse, human, frog, quail, sheep, cow, dog, rabbit, hamster and monkey [Izsvák, 2000]. *SB* was able to increase the frequency of transgene integration in all of these cell lines, with the exception of the quail. These results indicated that SB is active in most vertebrate species what makes it a good tool for studying the biology of DNA-transposons.

1.3.2 The mechanism of *SB* transposition

In 1994, a model was suggested for the chemical steps of Tc1/*mariner* transposition based on analysis of Tc3 transposition and on the analogy with other transposable elements in prokaryotes and V(D)J recombination [van Luenen, 1994]. The model (shown in Abb. 3) is so called cut-and-paste mechanism. The transposition process can be essentially divided into four main steps: 1) binding of the transposase to the transposon DNA; 2) formation of a paired-end-complex; 3) excision from the donor site; 4) target capture and reintegration at a target site. *Sleeping Beauty* transposes via cut-and-paste mechanism. First, the *SB* transposase recognizes and binds four specific transposon sequences at the ends of the element (direct repeats, DR) what is followed by paired-end-complex (PEC) formation of transposase tetramer bringing two transposon ends in close proximity. Then, the transposase is able to make double-stranded DNA breaks at each end. The breaks occur by hydrolysis of the phosphodiester bond exactly at the 3'-end of the transposon and three nucleotides within the transposon at the 5'-end. (for some transposons the break at the 5'-end occurs two nucleotides from the terminus). Every DNA strand cleavage in transposition reaction is Mg^{2+}-dependent. The element is released from the donor leaving behind 3-nucleotide-long 3'-overhangs at the excision site [Luo, 1998]. The excision site will be repaired by the host factors generating 3 bp footprints. The OH group at the 3'-ends of the excised transposon attacks the phosphodiester bond that flanks the TA target at the 3'-ends. The cleavage of the phosphodiester bonds at the target site is carried out in one step with the fusion to the 3' ends of the incoming transposon. The integration occurs exclusively into TA dinucleotide [van Luenen, 1994], [Ivics, 1999], with a preference for certain sequences flanking the TA dinucleotide. However, the target site selection is determined primarily on the level of DNA structure, and not by specific base-pair interactions [Vigdal, 2002]. Upon insertion a five-nucleotide single-stranded gap at each transposon terminus is formed, that is thought to be repaired by host machinery.

1.3.3 Effects of host factors on *SB* transposition

For Tc1, Himar1 and Mos1 an *in vitro* assay has been developed with the transposase as the only protein suggesting that these reactions do not require presence of other proteins [Vos, 1996], [Lampe, 1996], [Tosi, 2000]. However, some host proteins may influence the efficiency of the transposition process. Indeed, for *SB* four host factors have been identified, HMGB1 [Zayed, 2003], Ku 70/80 [Izsvak, 2004] and Miz-1 [Walisko, 2006]. Miz-1 is a Myc-associated transcription factor [Peukert, 1997] and a potent cell-cycle regulator [Staller, 2001]. Through its association with Miz-1, the *SB* transposase down-regulates cyclin D1 expression prolonging this way the G_1 phase. A temporary G_1 arrest enhances transposition [Walisko, 2006]. Ku 70/80 is a repair factor involved in repairing DNA damage produced by transposon excision. HMGB1 is an abundant, non-histone, nuclear protein associated with eukaryotic chromatin [Bustin, 1999]. The protein interacts with DNA in a sequence-independent manner, bending the DNA [Thomas, 2001]. HMGB1 was shown to interact with *SB* and to enhance preferential binding of the transposase to the inner direct repeat within the transposon inverted repeats as well as promote bending of DNA fragments comprising the transposon inverted repeats [Zayed, 2003].

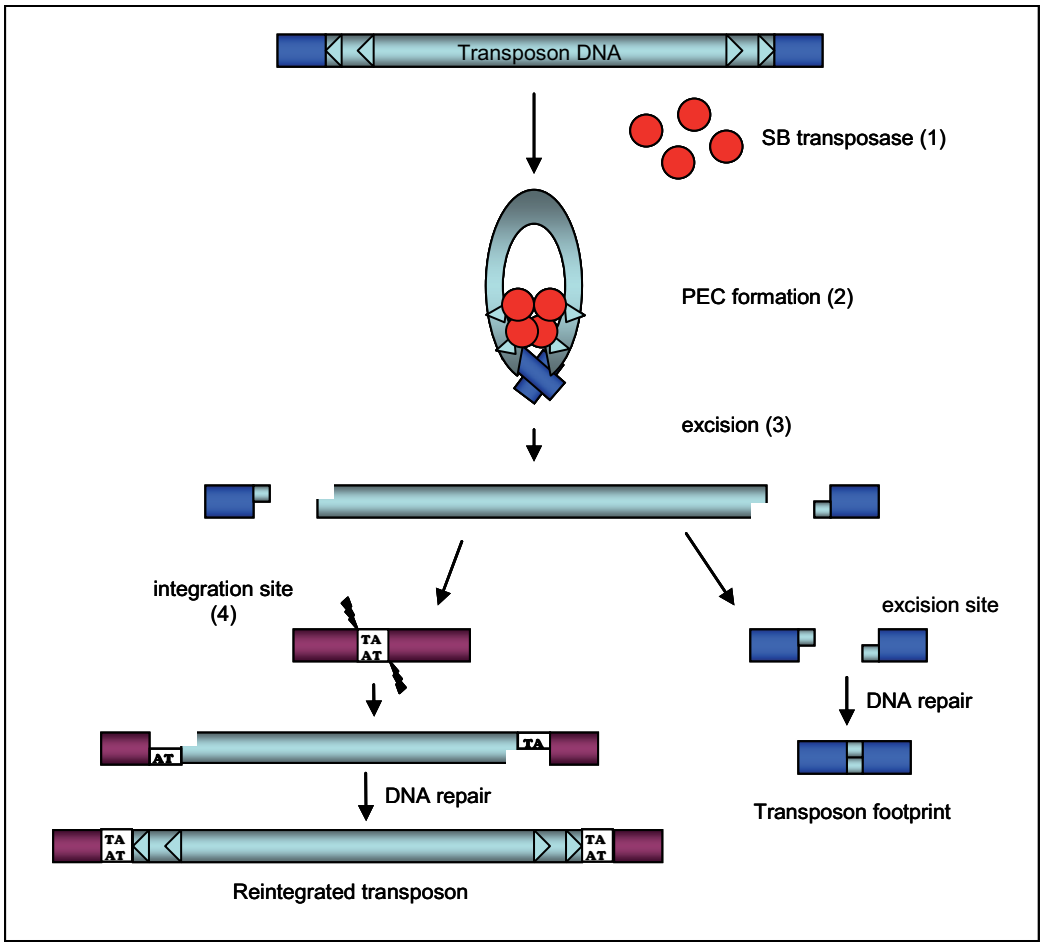

Abb. 3: Cut-and-paste mechanism of transposition. The transposition process can be essentially divided into four main steps: 1) binding of the transposase to the transposon DNA; 2) formation of a paired-end-complex (PEC); a tetramer of the transposase brings two transposon arms in close proximity; 3) excision from the donor site; 4) target capture and reintegration at a target site. Gaps at the excision and integration site are repaired by the host machinery.

1.3.4 Possible applications of the *SB* transposon system

Many new methods based on DNA transposons, retrotransposons, and even mobile introns have become practical *in vitro* as well as *in vivo*, affording new levels of specificity and efficiency. However, until recently, transposons were used as tools only in invertebrate systems, as no active vertebrate transposon was available. Resurrection of the *Sleeping Beauty* brought to life a promising tool for variety of applications. As previously mentioned, *SB* shows activity in many different animals making it a good tool for insertional mutagenesis. The system is being developed in such model organisms as mice and rats to facilitate the gene function analysis [Mátés, 2007]. Transposon based mutagenesis has several advantages; the

rate of DNA integration into chromosomes is high and the insertion site can be easily identified.
SB is a promising alternative to viral vectors for gene therapy. Vectors based on such viruses as retroviruses, lentiviruses, adeno-associated viruses have been developed and applied in gene therapy trials (reviewed in: [Kennedy, 1997], [Park, 2008], [Goins, 2008]. However, preferential integration of viruses into transcription units/genes [Schröder, 2002] together with strong enhancer/promoter elements in its long terminal repeats (LTRs) carry risks of undesired influence on expression of other, properly functioning, genes. Indeed, some of the patients with X-linked severe combined immunodeficiency (X-SCID) cured by gene therapeutic approach with retroviral vector developed leukemia upon insertion of the vector in proximity to T-cell oncogene [Hacein-Bey-Abina, 2003]. In addition, there is always the fear that the viral vector, once inside the patient, may recover its ability to cause disease. *SB* system might be a safer alternative to viral vectors as it shows fairly random integration profile [Vigdal, 2002], integration target site selection for retroviruses and transposable elements reviewed in [Wu, 2004]. Moreover, *SB* exhibits benign enhancer/promoter activities that can be reduced by incorporation of insulators [Walisko, 2008]. Importantly, the transgene carried by the *SB* transposon integrates at high rate and shows long-term expression. Additionally, unpublished data from our laboratory indicate activity of the *SB* system not only in dividing but also in non-dividing cells.
The *Sleeping Beauty* transposon system has proven efficacy in the treatment of hemophilia [Yant, 2000], tyrosinemia type I [Montini, 2002], junctional epidermolysis bullosa [Ortiz-Urda, 2003] and type 1 diabetes [He, 2004] in mouse models for human genetic disorders. Moreover, methods have been developed to deliver *Sleeping Beauty* transposons to the lung, liver and tumors for treatments for cystic fibrosis, cardiovascular and metabolic diseases, and cancer (reviewed in [Essner, 2005]. In the United States, the first clinical trials for cancer gene therapy using *Sleeping Beauty* system have been launched last year [Williams, 2008].
To improve the *Sleeping Beauty* system successful approach was undertaken in our laboratory to create more active enzyme [Mátés, 2009]. Additionally, fusion proteins of *SB* transposase with specific DNA binding domains are being developed to achieve targeted gene insertion [Voigt, 2008]. Deep understanding of the mechanism of *SB* transposition and its interactions with host environment would also contribute to improving the system in order to use it as a gene therapy vector.

1.4 Regulation of transposition

Mobilization of transposons must be under tight regulation, in both its frequency and insertional specificity, to avoid the accumulation of genome rearrangements that could be deleterious to the host. The level of transposition activity is probably the result of a balance between the interests of the transposable elements and those of the hosts they inhabit. Different host- and transposon-derived regulatory

mechanisms have been developed to keep the transposition activity at the desired level. In most instances, transposase is the only requirement for transposition. Therefore, the control of transposition is often limited to the regulation of the expression of this protein. Regulation of transposition is specific for each element and includes regulation at the levels of transcription, differential splicing, translation, and protein-protein interactions.
In most cases, the expressed protein is already tempered as conserved sequence of the transposase does not translate into the most active protein. Mutagenizing the *Sleeping Beauty* in our laboratory led to creation of a transposase 100 times more active than the originally created transposase [Mátés, 2009], [Ivics, 1997]. Mutagenesis resulting in creation of hyperactive transposases was performed also for Himar1 [Lampe, 1999], Mos1 [Pledger, 2005] and Tn5 [Naumann, 2002].
Endogenous transposase promoters are generally weak and many are partially located in the terminal IRs. This would enable their autoregulation by the transposase binding. The transposase promoters can be also bound by repressor proteins. The Mu repressor is a transpositional inhibitor that can also act as a transcriptional repressor of the transposase gene [Krause, 1986]. Transposition of some transposable elements is regulated by DNA adenine methylation. The methylation of the promoter region of the maize *Ac/Ds* elements reduces the transcription of active elements almost completely [Fedoroff, 1983], [Kunze, 1996]. In the case of other maize transposon, En/Spm, methylation of two regions close to the transcription start site inactivates the promoter. Additionally, the methylation reduces the excision frequency even when transactivated by an active En/Spm element [Schläppi, 1993]. In series of bacterial elements, including IS10, Tn903 and Tn5, increase in transposase expression levels and transposition frequency was observed in methylation deficient *dam*$^-$ strains [Roberts, 1985], [Yin, 1988].
In bacterial element, IS*200*, two sets of internal inverted repeats generate RNA secondary structures providing two independent mechanisms for repression of transposase synthesis [Beuzón, 1999]. One set forms a hairpin that terminates read-through transcripts before they reach the IS*200* ORF. Any impinging transcript which escapes the terminator will encounter a step-loop element that occludes the ribosome-binding site for the transposase gene, thus functioning as translational inhibitor. This double expression repression significantly decreases the levels of the IS*200* transposase.
In nematodes, RNA interference (RNAi) has been found to be a major mechanism for transposon silencing (reviewed in [Vastenhouw, 2004]. RNAi is the post-transcriptional silencing of a gene. Protein components of the RNAi machinery process double-stranded RNA (dsRNA) into small interfering RNAs (siRNAs) that mediate degradation of the sequence-complementary transcript. The dsRNA substrate for transposon silencing may arise from a transcription through the whole transposon including both inverted repeats. The two halves of the resulting transcript are complementary and can fold into an RNA hairpin. Moreover, the read-through transcription can occur both in a sense and antisense direction enabling the formation of dsRNA over the full length of the transposon. There is a growing evidence that RNAi is one of the mechanisms regulating transposition

levels in plants [Slotkin, 2009], fungi (Nolan et al., 2005) and animals [Houwing, 2007] including human [Yang, 2006].
Translational control of transposition by frameshifting has been demonstrated for IS1 [Sekine, 1989], [Luthi, 1990] and for members of the IS3 family [Polard, 1991], [Chandler, 1993] but may also occur in several other transposable elements. The upstream frame appears to carry a DNA recognition domain whereas the downstream frame encodes the catalytic site. The product of the upstream frame alone acts as a modulator of activity, presumably by binding the IR sequences. Frameshifting assembles both domains into a single transposase protein responsible for the mobility of the element. The frameshifting frequency is thus critical in determining overall transposition activity and could be possibly influenced by host physiology, thus coupling transposition activity to the state of the host cell.
En/Spm encodes a series of proteins that arise by alternative splicing. Two proteins, TnpA and TnpD, have been implicated in the transposition process. Other, such as TnpB and TnpC, are not essential but appear to play a role in regulating transposition [Masson, 1991]. Their structure suggests that they may act as negative regulators of the Spm transposase function. Interestingly, TnpA is a unique regulatory protein, as it is able to activate the inactive, methylated Spm promoter and to repress the active unmethylated promoter [Schläppi, 1994]. For the repression a DNA-binding and dimerization domains are necessary, whereas for the activation an additional C-terminal sequence of TnpA is required.
Shorter transposase versions may arise from translation from different start codons or from proteolytic cleavage of the expressed protein. If such proteins retain DNA-binding ability but the catalytic domain is disrupted, they can compete for DNA-binding with full-length transposase but will not form proper complexes. For several transposases, the N-terminal part of the protein exhibits higher affinity for the transposon ends than the entire transposase molecule. In such case, releasing the N-terminal DNA-binding domain would result in significant competitive inhibition. (Additionally, this fact suggests that the C-terminal end may in some way mask the strong DNA binding activity of the N-terminal domain.).
Another transposon-derived regulatory mechanism is overproduction inhibition (OPI). High levels of the transposase strongly decrease the overall transposition activity. The phenomenon has been observed for bacterial, plant and vertebrate transposons [Weinreich, 1994], [Kunze, 1993], [Geurts, 2003]. The mechanism is not completely understood, but based on studies on *Mos1 mariner* element it has been suggested that OPI results from posttranslational interactions between wild-type transposase molecules [Lohe, 1996]. The transposase monomers, when present in excess concentration, form inactive or weakly active oligomers. The OPI was also observed for *Sleeping Beauty* [Zayed, 2004].
As described above, there are many different mechanisms keeping the transposition at low level. However, to avoid genome rearrangements, not only the quantity of the “jumping” but also the quality of the process matters. The cut-and-paste transposition mechanism can be divided into four main steps: 1) binding of the

transposase to the transposon DNA; 2) formation of a paired-end-complex; 3) excision from the donor site; 4) target capture and reintegration at a target site (Abb. 3). Imprecision during complex formation or excision could lead to cleavage within the adjacent genomic DNA and taking up additional sequences or leaving part of the transposon DNA behind. To ensure precise transposition process, regulatory "checkpoints" exist, at which certain conditions have to be met for the reaction to proceed.

Once the transposase recognized its specific DNA ends, the complex can start to form. For some transposons the complex formation has been studied very deeply and there is evidence that it is a multistep complex reaction. Proteins of the DDE family require for the reaction the presence of Mg^{2+} ions for catalytic activity. Substitution for Ca^{2+} ions allows the complex formation but no catalytic reactions. It has been reported for *mariner* that *in vitro* transposition took place in presence of Mn^{2+} but the efficiency of the reaction was changed (15-fold lower than with Mg^{2+}, [Tosi, 2000]; 10-fold higher than with Mg^{2+}, [Zhang, 2001]. Interestingly, only half of the insertions took place at TA dinucleotide, whereas the rest was random. Similar observations were made for Tn10 [Junop, 1997]. The Drosophila P element requires for the process not only the presence of the divalent metal ions but, uniquely, also GTP [Kaufman, 1992]. First, the transposase binds to one of the two P element ends, in the presence or absence of GTP. However, in the absence of GTP, such complex remains stable but can not proceed to synapsis [Tang, 2005].

Biochemical studies of Tn5 transposase demonstrate that N-terminal DNA-binding domain and C-terminal catalytical domain inhibit each other's activities. No monomeric DNA-protein complexes have been detected when the full-length Tn5 transposase was used, whereas C-terminal deletion proteins could easily form complexes [Weinreich, 1994], [York, 1996]. These results are supported by structural analysis [Davies, 2000] and suggest that the full-length transposase must undergo a conformational change to relieve its activities. Importantly, the paired-end-complex formation, where two transposon ends are brought in close proximity, is a prerequisite for the catalytic activities, as Tn5 transposase acts in *trans*. It means that the transposase bound to one transposon end process the cleavage of the second transposon end [Naumann, 2000]. If proper complex formation fails, no excision would take place. This feature is shared by Mu [Savilahti, 1996] and most probably by other bacterial transposases. In contrary, *mariner*-like element *Himar1* was reported to form active complexes on one transposon end *in vitro* [Lipkow, 2004]. *mariner* transposases exist in solution as monomers, but are bound to a single inverted repeat (IR) on each end of the transposon as a dimer [Augé-Gouillou, 2005]. On the other hand, Butler and colleagues argue that *mariners* might bind as monomers and then quickly recruit another monomer to form a very stable dimer [Butler, 2006]. Nevertheless, once the dimer is bound, it is able to interact with similar IR-dimer complexes forming a paired-end-complex. However, a dimer of the transposase bound to one transposon end was able to cleave it when supplied with Mg^{2+} ions [Lipkow, 2004].

For *Sleeping Beauty*, there are many unanswered questions about the paired-end-complex (PEC) formation. It is known that a tetramer of the transposase is present in the PEC, however, how the complex forms has not been revealed. In this work I tried to dissolve the cascade of single steps and "checkpoints" necessary to form proper PEC that could undergo catalytic reactions.

2 MATERIALS AND METHODS

2.1 Materials

2.1.1 Bacterial strains

Plasmids were maintained in:

- *E.coli* DH5α

Expression of the proteins was performed in:

- *E.coli* BL21(DE3) CodonPlus RIL
- *E.coli* BL21(DE3) Rosetta pLysS

In vitro transposition assay was selected in:

- E.coli DH10β

Bacterial strains were routinely grown at 37 °C in LB medium

Luria Bertani medium:	10 g bactotryptone
	5 g bacto-yeast
	10 g NaCl
	dissolved in H_2O until 1 liter

Medium was sterilized by autoclaving and before use supplemented with the appropriate antibiotic (15 μg/ml of kanamycin or 100 μg/ml of ampicilin).

2.1.2 Plasmids

Plasmids available in the lab stocks of Zs. Izsvak's and Z. Ivics' groups:

- **pET-21a/N57 and pET-21a/58-123** expressing hexahistidine-tagged subdomains of the SB DNA-binding domain, PAI and RED, respectively [Izsvak, 2002];
- **MBP-SB**, plasmid expressing SB fused to maltose-binding protein [Zayed, 2003];
- **pET28/HMGB1** expressing hexahistidine-tagged version of HMGB1 was kindly provided by M. Bianchi, Milan, Italy (described in [Aidinis, 1999];
- **pCMV-SB10** [Ivics, 1997] and **pCMV-SBD3**, a catalytic mutant (E278D) of SB were used for expression of the SB transposase in HeLa cells;
- **pCMVβ,** plasmid expressing β-galactosidase used as a negative control in *in vivo* assays
- transposon donor construct **pT/neo** [Ivics, 1997];
- **pT/out**, a version of pT/neo containing in the right IR only the sequence of outer DR (constructed by Zoltan Ivics)
- **pCMV-Hsmar1-Ra** expressing an active version of Hsmar1 transposase [Miskey, 2007];

- **pHsmar1-neo** and **pHsma1-neoleft** lacking the right IR [Miskey, 2007];

Plasmids cloned during the project:

- **pT/in**, a version of pT/neo containing in the right IR only the sequence of inner DR followed by the TA dinucleotide;
- **5ir** and **5or**, pT/neo constructs with mutations introduced into the RED recognition motif of inner and outer DR of the right IR, respectively;
- **pTzeo/ss** – a pACYC184 plasmid with *rpsL* (streptomycin-sensitive allele) and *sacB* (sucrose sensitive gene) inserted as a *BglII-EagI* cassette and with a transposon DNA carrying zeocin resistance cloned into *EcoRI* and *NruI* sites (disrupting the Cm^R gene)
- **STREP-JAZZ-SB** was made by PCR amplification using three primers adding the tags N-terminally to the SB. The PCR product was then cut with *BamHI* and cloned into the *BamHI* site of the pIRES1*hyg* (CLONTECH) vector yielding a eukaryotic expression vector and into *NheI* site of pASK-IBA13plus (IBA BioTAGnology) yielding a bacterial expression vector;

2.1.3 Chemicals

alfa-^{32}dATP and alfa-^{32}dCTP	PerkinElmer
Agarose	MP Biomedicals
Amylose resin	NEB
Anhydrotetracycline (AHT)	IBA BioTAGnology
Antibiotic-antimycotic	Gibco/Invitrogen
Antibiotics (ampicilin, kanamycin)	Serva
Ammonium persulfate (APS)	Sigma
beta-mercaptoethanol	Sigma
bis(sulfosuccinimidyl) suberate (BS^3)	Pierce
Bovine Serum Albumin (BSA)	Sigma
bromophenol blue sodium salt	Sigma
COMPLETE Mini Tablets	Roche Diagnostics
Coomassie Brilliant blue R-250	Serva
dimethylsulphoxide (DMSO)	Sigma
dNTPs (100mM each)	GE Healthcare
dithiothreitol (DTT)	Sigma
Dulbecco's modified Eagle's medium (DMEM)	Gibco/Invitrogen
EDTA	Serva
ethanol	Roth
ethidiumbromide	Merck
Fetal Calf Serum (FCS)	PAA
formaldehyd 37%	Roth
GeneRuler 1 kb Ladder Plus	Fermentas
geneticin (G418)	Biochrom
glycerol 86%	Roth

glycine	Roth
HEPES	Sigma
imidazole	Sigma
isopropyl-beta-D-thiogalactopyranoside (IPTG)	Sigma
jetPEI RGD	Polyplus Transfections
maltose	Fluka
methanol	Roth
Ni-NTA agarose	Qiagen
N,N,N',N'-Tetramethylethylenediamine (TEMED)	Sigma
Nonfat dried milk powder	AppliChem
Nonidet P-40 (NP-40)	Fluka
PEI 25 kDa (linear polyethylenimine)	Polysciences
Phosphate buffered saline (PBS)	PAA
potassium acetate	Sigma
potassium chloride (KCl)	Merck
Precision Plus Protein Standards All Blue	Biorad
2-Propanol	Roth
ProtoGel 30%Acrylamide:0.8%Bis-Acrylamide	National Diagnostics
Roti-Phenol/Chloroform/Isoamylalcohol	Roth
sodium chloride (NaCl)	Fluka
sodium dodecyl sulfate (SDS)	Serva
Trizma	Sigma
Trypsin-EDTA, 5%	Gibco/Invitrogen
Tween-20	Merck
oligonucleotides for cloning	Biotez, Berlin

2.1.4 Kits

QIAGEN Plasmid Mini and Midi Kits	QIAGEN
QIAquick Gel Extraction Kit	QIAGEN
QIAquick PCR Purification Kit	QIAGEN
MicroSpin G-25 Columns	GE Healthcare
Ni-NTA Spin Columns	QIAGEN
Strep-Tactin Spin Column Kit	IBA Biotechnology
ECL Plus Western Blotting Detection System	GE Healthcare
pGEM-T Vector System	Promega

2.1.5 Enzymes and antibodies

Restriction enzymes	NEB, Amersham, Fermentas
Taq polymerase	Inviteq
Pfu polymerase	Stratagene
T4 DNA ligase	Fermentas
Shrimp alkaline phosphatase	Fermentas
Klenow fragment	Fermentas
DNase I	Invitrogen

Proteinase K	Invitrogen

Primary antibodies:

- Monoclonal Anti-Sleeping Beauty Transposase Antibody (R&D Systems) 1:500
- Polyclonal Anti-Sleeping Beauty Transposase Antibody (R&D Systems) 1:1000

Secondary antibodies:

- Goat Anti-Mouse IgG, Peroxidase Conjugated (Pierce Biotechnology) 1:5000
- Mouse Anti-Goat IgG, Peroxidase Conjugated (Pierce Biotechnology) 1:2000

2.1.6 Buffers

Buffers for electrophoresis of DNA	
TBE buffer	90 mM Tris-Borate, pH 8.3 1 mM EDTA
Loading buffer (5x concentrated)	50% glycerol bromophenol blue
Buffers for electrophoresis of proteins	
Running buffer	25 mM Tris, pH 8.3 250 mM glycine 0.1% SDS
Laemli buffer (4x concentrated)	0.25 M Tris, pH 6.8 5% SDS 10% β-mercaptoethanol 40% glycerol bromophenol blue
Coomassie Blue staining solution	0.2% Coomassie Brilliant Blue R-250 50% methanol 10% acetic acid
Western blotting	
Transfer buffer	25 mM Tris, pH 8.3 192 mM glycine 20% methanol
TTBS buffer	50 mM Tris, pH 7.5 150 mM NaCl 0.1% Tween 20
Electromobility Shift Assay	
Running buffer	25 mM Tris, pH 8.3 190 mM glycine 1 mM EDTA
Cell culture	
Fixing solution	10 % formaldehyde in PBS

Staining solution	methylene blue in PBS

2.1.7 Apparatus

ABI Prism 3100 DNA Sequencer	Applied Biosystems
Agigateur Top-Mix 11118	Bioblock Scientific
BioDocAnalyze	Biometra
Centrifuge 4K15C	Sigma
Centrifuges 5415D, 5417R, 5810R	Eppendorf
Fluorescent Image Analyzer FLA-3000	Fujifilm Life Science
Gelelectrophoresis, Mini-Sub Cell GT Gel System	Biorad
Gene Pulser II	Biorad
Image eraser	Raytest Isotopenmeßgeräte
Incubators for cell lines	Binder
Incubators for bacteria	Heraeus
Microscope Axiovert 25	Zeiss
Nanodrop	Peqlab
Optimax X-Ray Film Processor	Protec Medizintechnik GmbH
PCR Cycler Peltier Thermal Cycler-200/225	MJ Research
pH-meter	VWR International
Pipetboy	Brand
Pipettes, Pipetman	Gilson
Power supply electrophoresis	Biorad
SDS-PAGE, Mini-Protean	Biorad
Shaker, MTS4	Ika
Spectrophotometer, ND-1000	Peqlab
Thermomixer compact	Eppendorf
Vertical Gel Electrophoresis Apparatus, V15.17	Life Technologies
Waterbath	Julabo

2.2 Methods

2.2.1 Protein purifications

- *E.coli* BL21(DE3) CodonPlus RIL strain containing plasmid expressing MBP-SB fusion protein was grown to OD_{600} ~0.5 and after induction with 0.3 mM IPTG the growth continued for 2h at 37ºC. Purification was done as described in [Zayed, 2003] with small modifications. Bacterial pellet was resuspended in CB buffer (20 mM Tris pH 7.4, 200 mM NaCl, 1 mM EDTA and 1 mM DDT) containing protease inhibitors, DNase I and PEI and sonicated 6 times 15 seconds with 1 minute breaks. The pellet was resuspended in CB with 750 mM NaCl and centrifuged. The supernatant was loaded onto amylase resin equili-

brated in CB buffer. After washing steps, the fusion protein was eluted with CB buffer containing 10 mM maltose.

- Vector expressing STREP-SB was transformed into *E.coli* BL21(DE3) CodonPlus RIL and the bacterial culture was grown to OD_{600} ~0.6. After induction with 20 µg AHT the growth continued for 3h at 30ºC. Harvested bacteria was then resuspended in lysis buffer (50 mM NaH_2PO_4 pH 7.5, 500 mM NaCl) containing 1 COMPLETE Mini Tablet and sonicated in cycles 6 times 10 seconds separated with one minute break. After 20' centrifugation at 14 krpm the supernatant and the pellet were resuspended in Laemli buffer and analyzed on the SDS-PAGE gel.
- *E.coli* BL21(DE3) CodonPlus RIL strain was induced at OD_{600} ~0.6 with 20 µg AHT to express STREP-JAZZ-SB. After 4h growth at 30ºC the bacteria was centrifuged and the pellet resuspended in L-buffer (25 mM Tris 7.5, 250 mM NaCl and 1 mM EDTA) containing COMPLETE Mini Tablet and DNase I and sonicated (6 times 15 seconds with 1 minute intervals). The supernatant was loaded on equilibrated *Strep*-Tactin Column which was then washed with buffer W and STREP-JAZZ-SB was eluted with buffer E (both included in the kit).
- Expression and purification of His-tagged PAI and RED subdomains were conducted as described in [Izsvak, 2002]. Briefly, *E.coli* BL21(DE3) were induced at OD_{600} ~0.5 by the addition of IPTG to a final concentration of 0.4 mM and grown for 3h at 30ºC. Pellet was sonicated in H-buffer (25 mM HEPES pH 7.6, 1 M NaCl, 15% glycerol, 0.25% Tween20, 2 mM β-mercaptoethanol and a COMPLETE Mini Tablet) and the supernatant obtained after centrifugation was loaded on a Ni-NTA slurry equilibrated in H-buffer. After 2h incubation on a rotor at 4ºC the beads were washed with H-buffer containing 30 mM imidazole and bound proteins were eluted with H-buffer with 300 mM imidazole.
- The His-tagged HMGB1 was expressed in *E.coli* BL21(DE3) upon addition of 0.8 mM IPTG at OD_{600} ~0.6 and growth at 30^0C for 4h. The bacterial pellet was resuspended in lysis buffer (50 mM NaH_2PO_4 pH 7.8, 300 mM NaCl and 10 mM imidazole) containing 1 COMPLETE Mini Tablet and sonicated. Purification was done on Ni-NTA Spin Columns according to the manufacturer's protocol.

2.2.2 SDS-PAGE and Western Blot Analysis

Purifications steps of the tagged full-length transposases were checked on the 12.5% polyacrylamide gels. Purifications steps of the tagged subdomains, His-PAI and His-RED, as well as His-HMGB1 were separated on 15% polyacrylamide gels. The gels were run at 150 V, stained in Coomassie Blue solution for 30 minutes and distained incubating several times in warmed-up water.

To perform Western blot analysis, proteins from the gel were transferred onto nitrocellulose membrane (Amersham Biosciences) at 80 V for 1 hour 15 minutes in cooled transfer buffer. Membranes were blocked with 5% milk powder in TTBS for one hour at room temperature or overnight at 4°C. Then, the membranes were incubated with primary antibodies for two hours at room temperature or overnight at 4°C. After washing three times for 5 minutes with TTBS, the membranes were incubated with secondary antibodies for two hours at RT. The membranes were washed 5 times for 5 minutes with TTBS and the bound antibodies were visualized by chemiluminescence (ECL Plus Western Blotting Detection System).

2.2.3 Electromobility Shift Assay (EMSA)

Double-stranded oligonucleotides corresponding to the inner or outer DR were end-labeled using [alfa-^{32}P]dCTP and Klenow fragment. The DNA probe containing the left IR was cut out from the pT/neo using *EcoRI* restriction enzyme and end-labeled with [alfa-^{32}P]dATP using Klenow fragment. After the Klenow reaction labeled DNA was purified from nucleotides and protein using MicroSpin G-25 Columns as described by the manufacturer.
Binding reactions were performed in 20 mM HEPES (pH 7.5), 0.1 mM EDTA, 1 mM DTT in a total volume of 10 µl. 20000-50000 cpm labeled DNA probe and various concentrations of the proteins (as noted in the figures) were added and incubated 10 min on ice. After addition of 3 µl of loading dye (containing 50% glycerol and bromophenol blue) the samples were loaded onto a 4% or 6% polyacrylamide gel. The electrophoresis was carried out in Tris-glycine buffer pH 8.3 at 25mA for 2-3 hours. Then, the gels were dried for 45 minutes using the gel dryer from BIO-RAD. After overnight exposure the gels were scanned with Fujifilm FLA-3000 and analyzed with AIDA program.

2.2.4 Chemical crosslinking

Reactions were performed using the bis(sulfosuccinimidyl) suberrate (BS^3) according to manufacturer's recommendations. Proteins (3µM) were incubated on ice in 20 mM HEPES pH 7.5, 5 mM $MgCl_2$, 100 mM NaCl and 2.5 mM BS^3 in a final volume of 15 µl. After 2 hours the reactions were stopped by adding Tris pH 7.5 to a final concentration of 50 mM and incubating 10 min at room temperature. Then the Laemli buffer was added and samples were loaded on 15% SDS-PAGE and analyzed by Western blotting using polyclonal anti-SB antibody and secondary anti-goat IgG.

2.2.5 Cell culture and *in vivo* assays

HeLa cells were grown in DMEM supplemented with 10% Fecal Calf Serum Gold and 1% antimycotic antibiotic. One day prior transfection cells were seeded onto six-well plates. Cells were transfected with Qiagen purified DNA using jet-PEI RGD transfection reagent or 25kDa linear PEI in 150 mM NaCl. Two days

posttransfection half of the cells was harvested for excision assay and half was plated out on 10 cm dishes for selection using 1 mg/ml G418. After 3 weeks of selection, colonies were fixed with 10% formaldehyde and stained for 1 hour with methylene blue. After drying the plates were photographed and the colonies counted [Ivics, 1997].
To perform the excision assay plasmid DNA was isolated from harvested cells using the QIAprep Spin Miniprep protocol with small modifications. That is, instead of the P2 buffer, 1.2% SDS and 0.1 µg/µl Proteinase K was added. DNA was eluted in 50µl Elution Buffer and 2 µl was used as template for PCR. The PCR was performed with Taq polymerase and 10 pmol of primers aligning to the pUC19 backbone sequence in order to obtain the excision site.

PCR program: 94ºC - 94ºC – 65ºC – 72ºC – 72ºC – 4ºC
5’ - 30” - 30” - 30” - 5’ - ∝
30 cycles

The PCR product was diluted 1:100 and 1 µl was subjected to another round of PCR with nested primers. The products were visualized on the 1.2% agarose gel. To normalize the PCR conditions for excision assay, a PCR for ampiciline gene was performed using 15 cycles of the PCR program.

2.2.6 Analysis of excision and insertion sites

In order to analyze the excision sites obtained PCR products were cut out from the gel and the DNA was extracted with QIAquick Gel Extraction Kit. Then, the DNA was cloned into pGEM-T vector system, the positive colonies were picked and the sites were sequenced using the T7 primer.
To analyze the integration profile of the T2 construct a transposition assay in HeLa cells was performed. After 2 weeks of G418 selection the resistant clones were picked and transferred to 6-well plates. They were kept in culture for several days to grow in size and when the wells were full, cells were washed with PBS and 1 ml of lysis buffer (10 mM Tris pH 8.5, 5 mM EDTA, 0.5% SDS, 200 mM NaCl and 100 µg/ml Proteinase K) was added onto the cells. After overnight incubation at 37ºC the lysed cells were transferred to eppendorf tubes and the DNA precipitated. To the pellet 200 µl TE with RNase was added and incubated overnight at 55ºC with moderate shaking to dissolve the genomic DNA. 1 µg of the DNA was digested with BglII and BclI restriction enzymes, precipitated and ligated overnight at 16ºC in a large volume to allow the DNA fragments to circularize. After precipitation the DNA samples were resuspended in 30 µl TE buffer and 2 µl was used as template for PCR reactions. Nested PCRs were conducted with primers aligning to the neomycin resistance gene and inverted repeat sequences. Obtained products were extracted from the gel and sequenced.

2.2.7 *In vitro* transposition assay

200 ng donor (plasmid containing transposon sequence) DNA was incubated together with 50 ng of purified transposase fusion protein in buffer containing 25 mM HEPES, 100 mM KCl, 5 mM $MgCl_2$ and 1 mM DTT. For some reactions the buffer was enriched with 10 % glycerol, 10 % DMSO and/or 0.02 % NP40. Additionally, different ratios of DNA and protein were used. Some reactions were supplied with 50 ng of purified His-HMGB1 and/or 10 ng of BSA. After 30 minutes incubation at 4ºC 2 µg target DNA was added and incubated 3 hours or overnight at 30ºC. The reactions were stopped by adding 5 % SDS, 2 mg/ml proteinase K and 50 mM EDTA in 25 mM HEPES and incubating at 55ºC for 1 hour. After phenol-chlorophorm extraction the DNA was precipitated (0.6 volume of isopropanol, 0.1 volume of 2.5 M KoAc and 1 µl glycogen), washed twice with 70 % ethanol and the pellets were resuspended in 10 µl H_20 and electroporated into DH10B cells. To recover the transposition events the cells were selected on LB agar containing ampicilin, zeocin, streptomycin and sucrose.

3 AIM OF THE PROJECT

Transposons are widely spread in nature. The transposition has to be controlled in order to keep the process at low level. Uncontrolled transposition could be deleterious to the host and thus to the transposon itself. That is why both, transposon and the host, developed mechanisms regulating transposition reaction.
Until the resurrection of the *Sleeping Beauty* [Ivics, 1997] no active vertebrate transposon was available. The *SB* system is now being developed as a tool for insertional mutagenesis as well as a vector for gene therapy. Studies of *SB* transposition *in vitro* and regulatory mechanisms of the process would provide information essential for the optimization of transposition recoveries *in vivo*. Therefore, the aim of the project was to understand the single steps of the transposition reaction and to establish an *in vitro* assay for *Sleeping Beauty* transposition. Insights into the process and knowledge about conditions required for each step of the reaction would help in future work and use of the transposon system. The *in vitro* assay would serve a quick and easy assay for testing the system and probe mechanisms affecting the regulation of transposition activity.

4 RESULTS PART I - Regulated complex assembly

4.1 IR/DR structure of Sleeping Beauty

Sleeping Beauty belongs to IR/DR subfamily of Tc1/*mariner* transposons. The IR/DR structure refers to two Inverted Repeats (IR) each having two Direct Repeats (DR) flanking an open reading frame coding for the transposase. The inverted repeat upstream of the gene is commonly called left IR and the inverted repeat downstream of the gene, right IR. Each IR is approximately 230bp long and has two 30bp long DRs serving as binding sites for the transposase. All four binding sites are necessary for the transposition process, as deletion of one of them has a strong negative effect on the transposition efficiency. Experiments performed in our laboratory (by Christopher D. Kaufman, Abb. 4) revealed that removing one binding site from the left IR abolishes transposition, whereas removing one or both binding sites from the right IR strongly decreases transposition. Additionally, the distance between the binding sites matters as removing this sequence abolished transposition [Izsvak, 2002]. Moreover, the position of the outer and inner binding sites can not be freely exchanged with each other [Cui, 2002]. Not only are the presence of all binding sites, but also the combination of the inter-DR spacing and geometry of the DRs important for the transposition process. In *SB* system all binding sites are occupied by the transposase proteins and the active nucleoprotein complex is composed of transposase tetramer bringing two transposon ends together [Izsvak, 2002]. Similar complex is formed for Tc3, another Tc1/*mariner* element with four binding sites. Interestingly however, deletion of inner binding sites of the Tc3 transposon DNA does not abolish its transposition [Colloms, 1994].
mariner-like elements form the second subfamily of the Tc1/*mariner* transposons. In contrast to *SB*, *mariners* have short IR with only one binding site each. Importantly, elements deprived of the right IR showed moderate transpositional activity. Formation of active complexes on single transposon end, without the second end in the complex, was reported for the *mariner*-like element, *Himar1* [Lipkow, 2004]. It means that a transposase dimer formed on one binding site is able to cleave the transposon end.

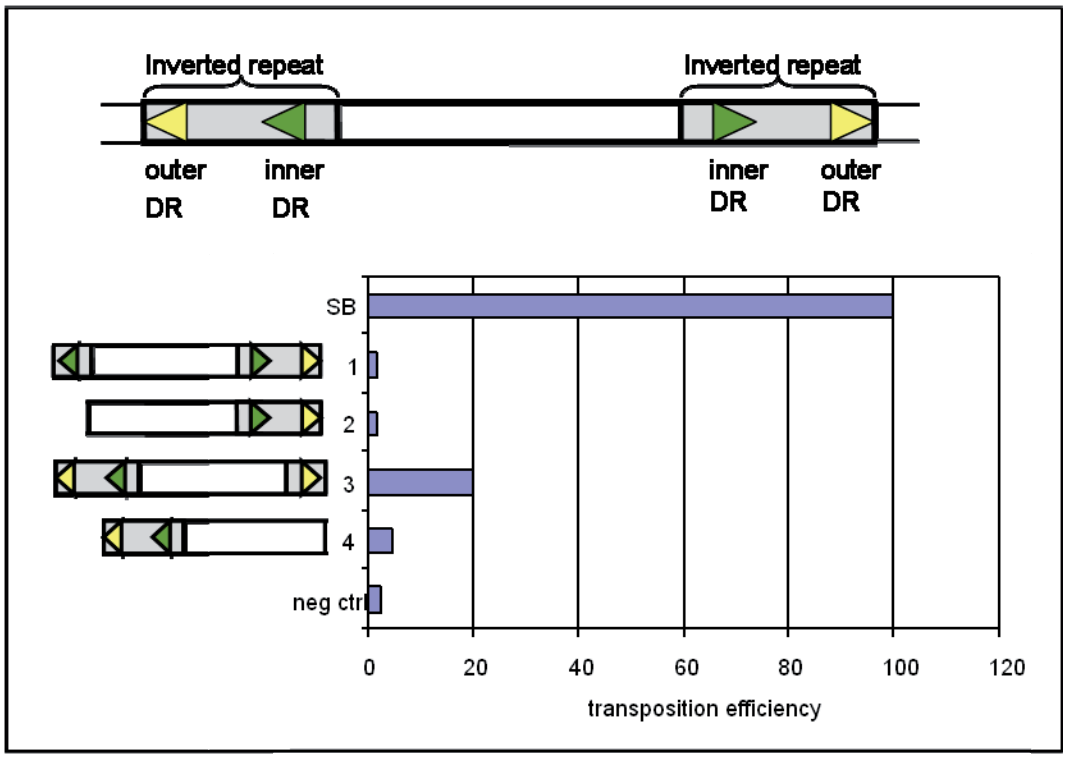

Abb. 4: *SB* requires all binding sites for proper transposition. On the top, the DNA structure of the *SB* transposon with yellow and green triangles referring to outer and inner binding sites (DR), respectively. Deletion of one (construct 1) or both (construct 2) direct repeats from the left inverted repeat completely abolishes transposition and deletion of one (construct 3) or both (construct 4) direct repeats from the right inverted repeat strongly decreases the transposition efficiency in transposition assays *in vivo*.

Since transposon integration events with *SB* construct having only one transposon end were detected (Abb. 4, construct 4), suggesting that *SB* transposase is as well able to form active complexes on one inverted repeat, I wanted to test and compare the frequencies of single-end cleavage of *Sleeping Beauty* and a *mariner*-like element. In order to do that, I subjected full-length (containing both inverted repeats) and left-arm (missing the right inverted repeat) transposons of *SB* and a *mariner*-like element, *Hsmar1* [Robertson, 1997], together with their cognate transposase-expressing helper plasmids to *in vivo* transposition assay [Ivics, 1997]. All transposons contain a neomycine resistance gene. Cells with transposition events can be then selected using G418 antibiotic. As negative controls, a plasmid expressing β-galactosidase (pCMVβ) was used instead of plasmids expressing active *SB* and *Hsmar1* transposases. Fischer and colleagues [Fischer, 2001] compared transposition efficiency of *SB* and several other Tc1/*mariner* elements in transposition assays in HeLa cells. Similarly to their results, I observed higher transposition efficiency for *SB* than for *Hsmar1* (Abb. 5A).

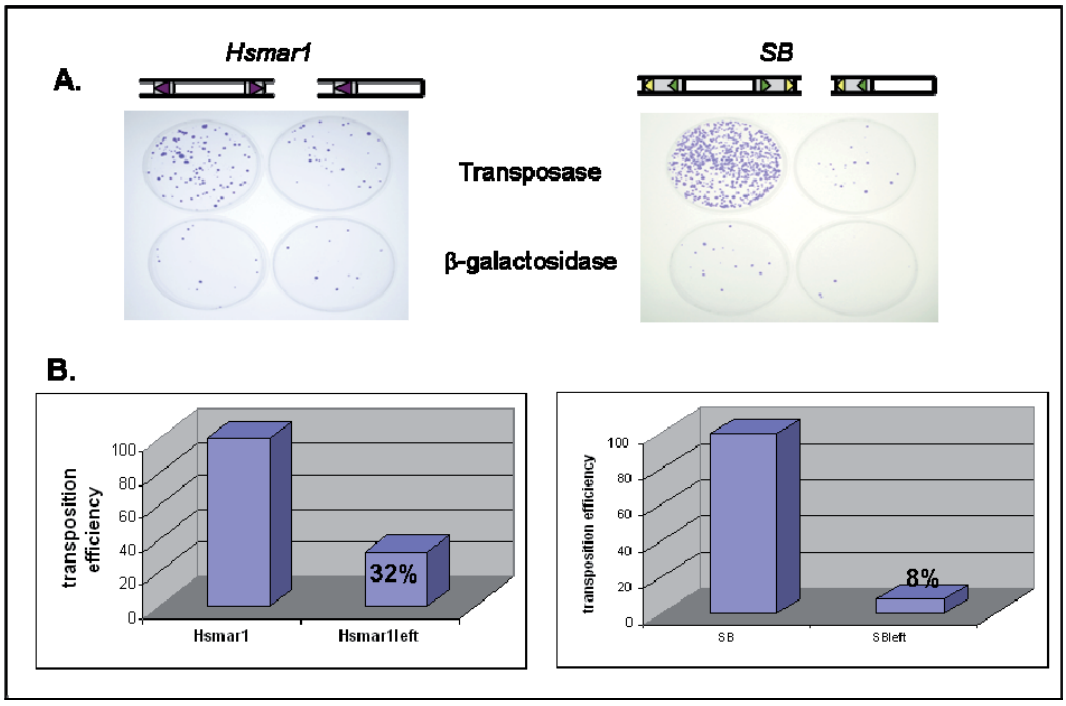

Abb. 5: SB is capable of single-end-cleavage. A. transposon constructs carrying a neo resistance gene were cotransfected with a plasmid expressing either transposase or beta-galactosidase. G418-resistant colonies were selected and counted. **B.** The transposition activity (indicated on the *z* axis) is the ratio of the number of resistant colonies obtained in the presence of the transposase over the resistant colony number obtained with β-galactosidase. The activity of the full-length transposons was set to 100%. The activity indicates an average of three independent transfections.

For all transposon constructs, full-length as well as left-arm, raise in colony numbers in comparison to negative controls was observed (Abb. 5A), indicating that both transposons are capable of single-end cleavage. However, the frequency of such events is significantly lower for *Sleeping Beauty* (Abb. 5B). As mentioned above, the main difference between *SB* and *Hsmar1* systems is the inverted repeat structure. Additionally, the N-terminal DNA-binding domain of the *SB* transposase has a more composite structure. The domain consists of two helix-turn-helix motifs, called PAI and RED [Izsvak, 2002] (Abb. 2A), while the transposase of the *Hsmar1* has only one HTH motif [Robertson, 1997]. Existence of two DNA-binding subdomains within *SB* transposase may indicate more complex protein-DNA interactions. I wondered what the function of the bipartite DNA binding domain is and performed experiments to learn more about the activities of the subdomains.

4.2 Functions of the bipartite DNA-binding domain of SB *transposase*

In order to find sequences bound strongly by the DNA-binding domain of the *SB* transposase, a PCR-based CASTing experiment was performed in our laboratory. The CASTing method was developed to select optimal binding sites for DNA-binding proteins [Wright, 1991]. Using footprinting data of *SB* transposase binding [Ivics, 1997], a 24 bp random oligonucleotide library was created and exposed to the DNA-binding domain of the transposase. Oligonucleotides selected in the CASTing experiment were sequenced and compared to the wild type binding sites. The binding sites for the *SB* transposase are very similar, but not identical, showing 3 nucleotides differences in sequence and 2 nucleotides differences in length (Abb. 2B). From footprinting data it was known that PAI recognizes the 3'-part of the binding site, whereas RED recognizes the 5'-part [Izsvak, 2002]. Surprisingly, all of the selected sites showed similarity only to the PAI recognition motif, whereas the RED recognition site was fairly random. Interestingly, when the PAI recognition sequences within the inner binding sites were changed to one of the high affinity binding sites (the RED recognition sites were left wild-type), no significant raise in transposition rate was observed in transposition assay in cells. Replacement of the PAI recognition motif within the outer binding sites led to the drop of the transposition frequency (experiment by Christopher D. Kaufman). These unexpected results show that stronger substrate binding does not lead to better transposition suggesting that enhancing one step of the reaction may disturb the following steps.
Additionally, the randomness of the RED recognition site would suggest that the RED is not involved in the primary substrate binding and only the PAI is responsible for the primary interaction to its DNA substrate. This result raises also the question, what is the function of the RED subdomain?
It has been shown that the N-terminal PAI subdomain is not only able to recognize and bind the *SB* DNA, but has also a protein-protein interaction interface [Izsvak, 2002]. To study the activities of the RED subdomain, I purified a His-tagged fusion of the RED and performed electrophoretic mobility shift assay (EMSA) and cross-linking experiment with it. EMSA is an electrophoresis technique used to detect and study protein-DNA interactions. When the protein is able to bind given DNA, the formed complex migrates slower in the gel resulting in appearance of an additional shifted band.
SB binding was previously tested using either the entire IR or oligonucleotides corresponding to inner and outer DRs [Zayed, 2003]. Subjecting the RED subdomain to EMSA together with inner binding site resulted in appearance of one shifted band (Abb. 6A, well 2). Although PAI and RED subdomains are of very similar molecular mass, the shift of the RED-DNA complex is visibly stronger (compare lane 2 and 3 in Abb. 6A). This slower migration could suggest that a dimer of the RED binds the oligos or that the RED dimerizes quickly after DNA binding. Titrating the RED, I did not observe bands that could correspond to RED monomer in complex with DNA (data not shown). Additionally, to confirm the

ability of the RED to oligomerize, a cross-linking with chemical crosslinker BS^3 was conducted. Clearly, RED dimers and tetramers were formed upon incubation with BS^3 (Abb. 6C).

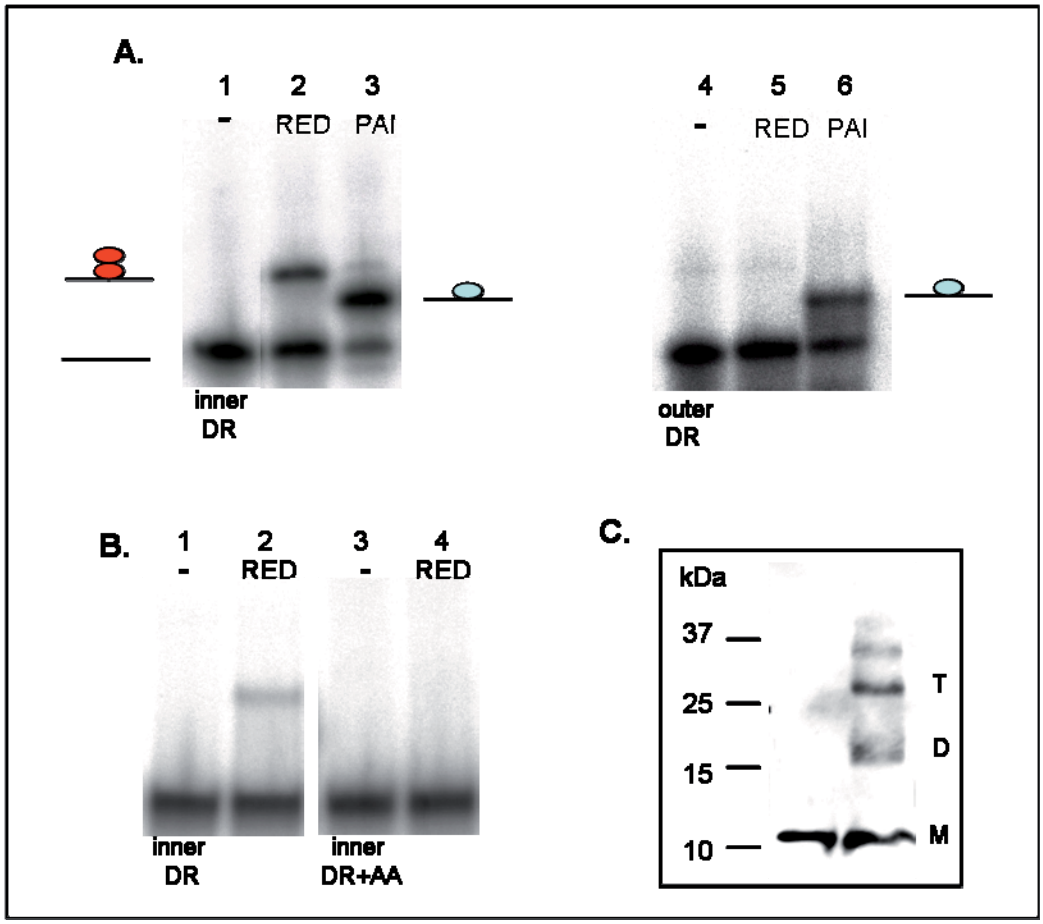

Abb. 6: RED subdomain has DNA-binding and protein-protein interaction activity. **A.** demonstration of the DNA binding ability of the RED subdomain in EMSA (lane 2). The oligonucleotides corresponding to DRs were incubated with or without the PAI and RED subdomains and separated on 6% polyacrylamide gel. The molecular composition of each complex is shown on the right and the left of the gels; **B.** binding assay testing the ability of the RED subdomain to bind the inner DR containing two additional adenosine nucleotides (inner DR+AA) (lane 4). **C.** multimerization of the RED subdomain in the presence of BS^3 crosslinker. Formed oligomers were detected in Western blot with polyclonal antibodies against SB. M, D and T stand for monomer, dimer and tetramer, respectively.

An interesting feature of the RED subdomain was revealed by the EMSA performed with the outer binding site. Even though RED was able to recognize and bind the inner DR, no binding to the outer DR was ever detected (Abb. 6A, lane 5). The binding sites differ from each other in two base pair in length and three nucleotides in sequence (Abb. 2B). I performed EMSA with inner DR containing

two additional adenosines present in the outer DR. The oligonucleotide (called inner DR+AA, Abb. 6B) is of the length of the outer DR, however, retaining the rest of the inner DR sequence. No binding to such modified DNA was observed (Abb. 6B, compare lanes 2 and 4). Binding affinity of the RED subdomain is very specific, restricted to the inner binding site of the SB transposon, as introduction of two nucleotides abolishes the binding. Important for distinction between the inner and outer binding site by the RED subdomain is the length of the two DRs. The PAI subdomain was able to bind both DRs (Abb. 6A, lane 3 and 6); however, the inner DR was bound stronger. The stronger affinity of both subdomains towards the inner binding site can suggest that the complex formation would start at the inner DR, distantly from the cleavage site.
As RED did not recognize the outer binding site, I wanted to check whether the sequence is of importance for the transposition reaction at all. In order to see this, I exchanged the RED recognition motif within the outer DR of right inverted repeat for one of the random oligonucleotides selected in the CASTing experiment (leaving the PAI motif wild-type). Additionally, the same sequence was introduced instead of the inner DR. Both constructs, 5ir and 5or, carrying the random site in the inner and outer DR, respectively, were subjected for a transposition assay *in vivo*, together with wild-type transposon and transposase as a positive control and mutant inactive transposase as a negative control. Although, drop in transposition efficiency was observed for both mutated transposons, it was more significant when the outer binding site was mutated (Abb. 7). Even though the RED subdomain did not show binding to the outer DR in EMSA experiments, data from transposition assays suggest that the RED-outer DR interaction must be important at some stage of the transposition reaction, as substitution for a random sequence strongly decreased the efficiency of the process.

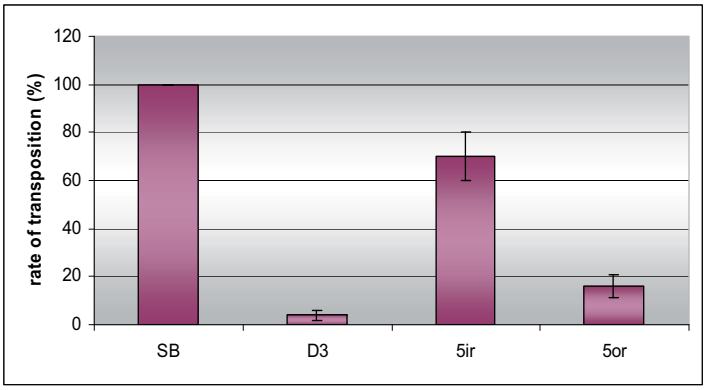

Abb. 7: RED recognition motif of the binding sites plays important role in the transposition. Substitution of the RED recognition motif decreased transposition efficiency in HeLa cells. Standard errors of the mean are from three independent transfections. D3 stands for inactive transposase mutant serving as a negative control.

4.3 paired-end complex (PEC) formation

The transposon-transposase complex formation is believed to be a multistep reaction with several DNA-protein and protein-protein interactions that take place in a regulated sequential manner. Previous EMSA experiments suggest that the primary binding starts at the inner DR, distantly from the cleavage site, and most probably upon binding the protein-protein interaction activity of the RED subdomain is released (Abb. 6). When excess of purified RED subdomain was supplied an additional band appeared in the binding experiment (Abb. 8A, lane 3 and 4) believed to be a dimer of the RED bound to two DNA molecules. To form PEC transposase has to interact with both ends of the transposon DNA. Appearance in the EMSA of bands corresponding to higher complexes suggested that RED subdomain might be responsible for catching the other arm of the transposon. An EMSA was performed to check which DR, inner or outer, is incorporated as next one into the paired-end-complex. First, the RED subdomain was incubated with the inner binding site and then increasing amounts of radioactively labeled inner or outer binding sites were added (Abb. 8B). Upon addition of the inner DR increase in higher complex formation was observed (Abb. 8B, lanes 2 to 4), whereas upon addition of the outer DR the complex was not formed (Abb. 8B, lanes 5 and 6). This result implies that the transposase after forming a dimer on one inner binding site catches the second inner binding site.
Incubation of increasing amounts of purified PAI subdomain with oligonucleotides corresponding to the binding sites never resulted in formation of higher order complexes. Only one band could be observed (data not shown). These results suggest that the PAI subdomain does not exhibit abilities to catch the second naked DNA.

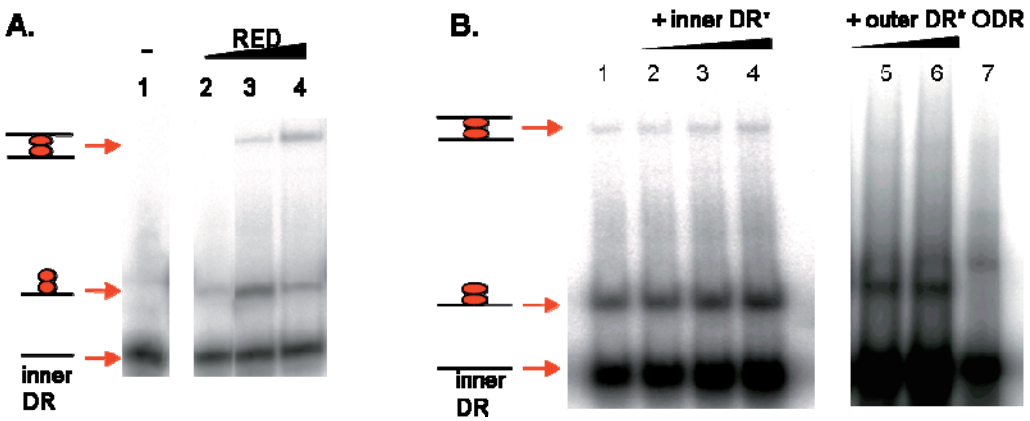

Abb. 8: RED catches the second transposon end. A. EMSA with increasing amounts of the purified RED subdomain. **B.** after the complexes are formed increasing amounts of labeled inner or outer DR are added to the reaction. Drawings on the left side of the gels show the molecular composition of each complex. ODR in lane 7 stands for outer DR alone.

It has been published that HMGB1 enhances binding of the *SB* transposase to the transposon IRs [Zayed, 2003]. HMGB1 is a nuclear protein binding the DNA in a sequence-independent manner and bending the DNA [Thomas, 2001]. Bending the DNA strand may bring two distant fragments in linear DNA in proximity. For transposons it would mean that HMGB1 brings closer the binding sites at the ends of the transposon DNA. For bacterial transposons presence of bending proteins, such as IHF and HU, in the complex formation has been reported [Krause, 1986], [Signon, 1995]. To test, whether HMGB1 influences DNA binding of the RED subdomain and the complex formation, the binding reactions were performed in presence and absence of HMGB1 and separated on the gel in electromobility shift assay (Abb. 9). Interestingly, when HMGB1 was present in the reaction the band corresponding to the paired-end complex was enhanced (compare lanes 2-4 and 5-7 in Abb. 9). In the EMSA oligonucleotides corresponding to the binding site were used so the formed complex consists of a dimer of RED binding two single DNA molecules and not one bent DNA. Nevertheless, HMGB1 clearly enhanced formation of the higher order DNA-protein complex. The HMGB1 protein itself is not stably included in the complex or its interaction is not as strong as no super-shift was observed and no shift when only HMGB1 was present (Abb. 9A, last lane). The mechanism of HMGB1 enhancement on DNA binding and formation of higher order complexes is not known.

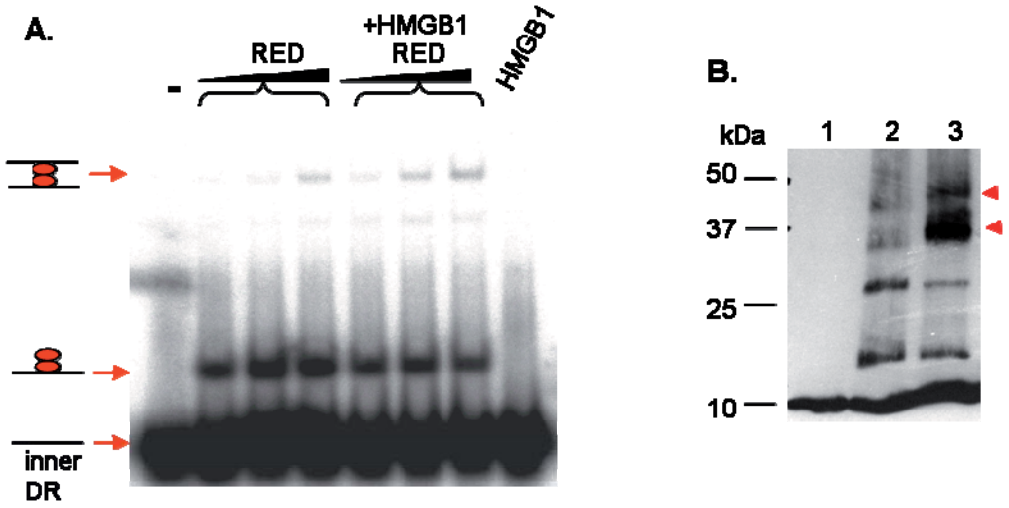

Abb. 9: HMGB1 enhances the RED binding to the DNA and catching of the second transposon end. A. electromobility shift assay of the RED subdomain in presence (lanes 5-7) and absence of HMGB1 (lanes 2-4). The bands correspond to complexes shown on the left side of the gel. **B.** chemical cross-linking of the RED with HMGB1 (lane 3). In lane 2 cross-linking of the RED alone used as a control.

HMGB1 was not only shown to enhance the *SB* binding to the transposon DNA but also to interact with the transposase [Zayed, 2003]. As I observed the influ-

ence of HMGB1 on RED binding activity, I checked the protein-protein interaction in the cross-linking experiment (Abb. 9B). When both proteins were incubated together in presence of the BS^3 two bands were detected (Abb. 9B lane 3, red triangles) that were not present without HMGB1 (Abb. 9B, lane 2). The size of approximately 37 and 45 kDa suggests complex of HMGB1 with a monomer and a dimer of the RED subdomain, respectively (29 kDa HMGB1 + 8 kDa RED).

4.4 Cleavage

The paired-end complex formation of *SB* transposon system starts at the inner binding site, distantly from cleavage site. My EMSA experiments results suggest that the next binding site incorporated into the complex is the inner DR of the second IR. A dimer of the transposase brings two ends of the transposon together. The dimer can be an active transposase form for related *mariner*-like elements. Even though *Himar1* PEC consists of transposase tetramer, a dimer bound to one transposon end was able to accomplish the catalytic activities [Lipkow, 2004]. We wondered whether the *SB* dimer bound to two inner DRs is able to process cleavage at the inner binding sites or in proximity to it. In order to test it, we planned to perform transposition and excision assays with truncated transposons with IRs ending with inner DRs. However, as deletion of outer binding site from the left inverted repeat results in complete abolition of the transposition reaction (Abb. 4, construct 1), the left IR had to be left unmodified. In the right IR only the sequence corresponding to the inner DR was present (Abb. 10A). At the DR-backbone border I introduced the TA dinucleotide to mimic the cleavage site. Additionally, the transposon carrying in the right IR only the outer binding site was included in the experiment (construct 3 shown in Abb. 4). Both constructs, called T/in and T/out, were transfected into HeLa cells together with a positive and negative control. The excision efficiency was tested in a PCR based assay recovering the excision site. Interestingly, excision product was detected for T/out, the transposon with outer DR, but not for T/in, the transposon with inner DR, (Abb. 10C, lanes 4 and 5). The excision PCR is not a quantitative method, nevertheless, as the assay was normalized amplifying the ampiciline gene present in the backbone (Abb. 10B) und thus determining the input DNA, the efficiency could be roughly estimated (Abb. 10D, blue bars). Even though the constructs differ only in one binding site, what actually comes down to 3 bp difference in sequence and 2 bp difference in length, they exhibit significantly different activity in excision suggesting existence of a mechanism preventing the cleavage at the inner DR.

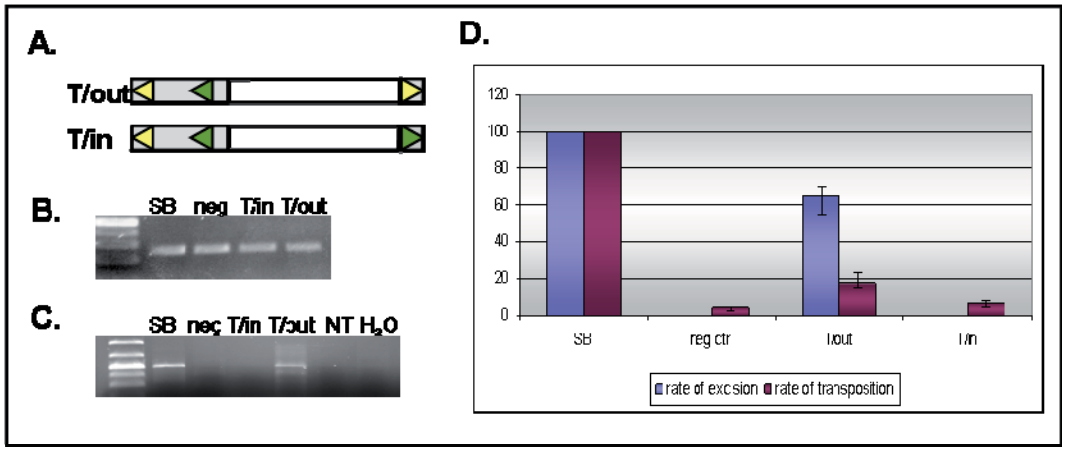

Abb. 10: cleavage is inhibited at the inner binding site. A. schematic representation of the T/out and T/in transposons, carrying in the right IR only outer or inner DR, respectively. **B.** the products of the PCR amplifying the ampicillin gene present in all vectors serving as the input control. **C.** the excision products detected in the PCR based excision assay. NT stands for non-transfected; neg stands for transfection with catalytically inactive transposase. **D.** excision and transposition efficiency of modified transposons. The bars represent the average of three independent transfections.

In transposition assay, both constructs showed significant decrease of the efficiency, around 80% for the T/out confirming previous results (Abb. 4) and up to 92% for the T/in (Abb. 10D, violet bars). Thus, the excision of the T/in took place, as low number of transposition events was formed in the colony based assay. However, the low level of excision was not detectable in the PCR based excision assay (Abb. 10C). The excision product of the T/out was sequenced to check the precision of the cleavage. Although canonical footprint was detected, most of the excision sites showed aberrancies. Additionally, individual G-418-resistant colonies were expanded to clones and the transposon integration sites were analyzed. Only proper insertion sites were found what may suggest that exclusively the transposons precisely excised from the donor site would undergo insertion into the new locus.

5 RESULTS PART II - Establishing an *in vitro* transposition assay

In vitro transposition assays have been established for some transposable elements such as yeast Ty1, bacterial Tn5 and ISY100, and insect *Mos1* [Devine, 1994], [Goryshin, 1998], [Tosi, 2000], [Feng, 2007]. The lack of higher number of *in vitro* assays can be due to the tight regulation of the process with strong requirements for proper conditions at each step and to dependence on the host factors. *mariners* seem not to require strict conditions as, as long as the transposase, donor (transposon) and target DNA, and divalent magnesium ions are supplied, they will jump [Tosi, 2000]. Studies of Tosi and Beverley provided information crucial for the design and application of new transposable elements for use in functional genetic analysis. The data obtained facilitates the development of the *in vivo* system. Establishing an *in vitro* transposition assay for the *Sleeping Beauty* would yield information about the requirements for optimal run of the process and additionally afford an assay to explore the regulatory mechanisms affecting the *SB* transposition. Collected data would help in development of the system for gene therapy and mutagenesis applications.
The activity of the *SB* system in a broad range of organisms, from fish, through amphibians and rodents to primates including human, suggests independence on the host factors [Izsvák, 2000]. However, the influence of host factors on *SB* transposition cannot be excluded. It is also possible that the host factors acting on the *SB* transposition are conserved through the vertebrate species.
Based on the *in vitro* system of the closely related *mariner* transposon, in the *SB in vitro* assay the purified substrates are incubated together and then the transposition products are recovered and electroporated into *E.coli* (Abb. 11). The transposon carries an antibiotic resistance gene between its inverted repeats resulting in a double-resistant target plasmid upon transposition. However, also bacteria that took up two separate plasmids may grow on double selection. To minimalize the background arising from electroporation of both plasmids, a donor plasmid with oriR6K (present also in the construct for *Mos1 in vitro* assay, [Tosi, 2000]) was used. oriR6K is origin of replication requiring the Pir protein for propagation. Virtually all common *E.coli* strains are *pir-* and thus, the donor plasmid containing the oriR6K should not be propagated in *E.coli* and only bacteria carrying transposition events can grow on double selection. However, the system turned out to be leaky as colonies resistant to both antibiotics did not result from transposition but from electroporation of both plasmids. In order to recover only transposition events another approach was undertaken introducing a different negative selection. A donor plasmid was cloned carrying an *rpsL* and *sacB* cassette, coding for streptomycin and sucrose sensitivity gene, respectively. Bacteria carrying this plasmid cannot grow in presence of streptomycin and 5% sucrose. Thus, to sur-

vive on double-selection they have to carry a target plasmid with transposon sequence resulted from the transposition event.

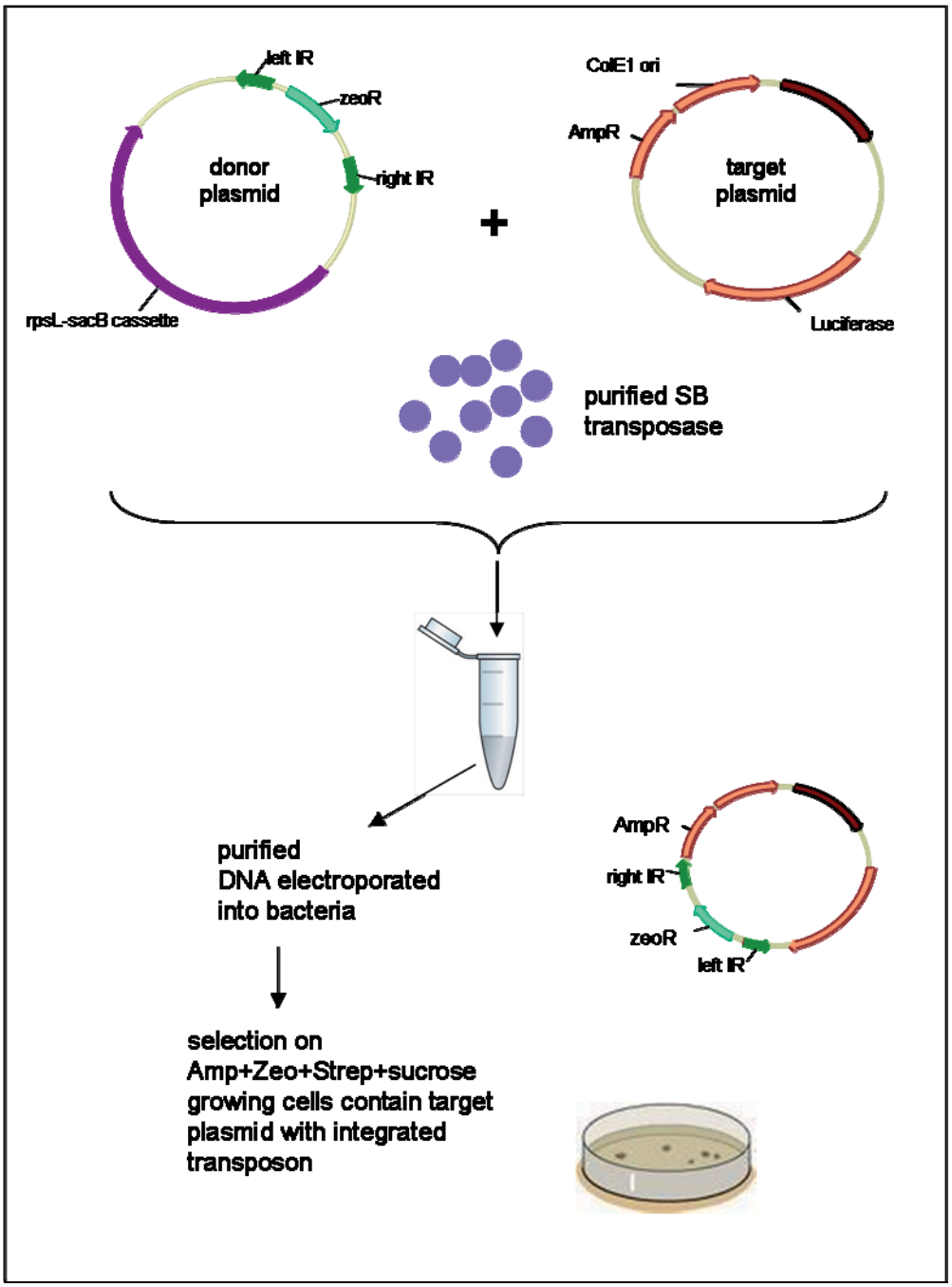

Abb. 11: ***In vitro*** **transposition assay.** Donor and target DNA are incubated together with purified transposase and after purification electroporated into bacteria and selected for transposition events.

Much effort has been previously undertaken to create an active purifiable SB-fusion protein. None of the C-terminal fusions showed any activity suggesting that the SB transposase does not sustain a tag fused to its C-terminus. Therefore, all of the proteins used in the experiments are N-terminal SB-fusions. First, in the assay shown in Abb. 11 a maltose-binding protein-SB fusion was tried (MBP-SB). MBP is a large tag increasing the stability and solubility of the fused protein. Indeed, MBP-SB fusion was easy to purify (data not shown). A prerequisite for transposition is the DNA-binding activity of the protein that can be easily tested in an electromobility shift assay. Therefore, the purified MBP-SB fusion protein subjected to the *in vitro* assay was checked for its ability to bind the *SB* DNA. I performed an EMSA using left IR as a probe and observed two shifted bands. One is believed to be a protein monomer bound to one binding site and the slower migrating band could correspond to two transposase monomers bound to two binding sites present in the inverted repeat (IR) or a dimer formed on one binding site (Abb. 12A, lane 2 and 3). To solve the molecular composition of the complexes, two IRs were used each missing one of the binding sites (Abb. 12B). If a dimer forms on the DNA, two shifted bands would occur corresponding to a monomer and a dimer bound to the present binding site, respectively. However, with mutated IRs only one band was observed (Abb. 12B) suggesting that the purified MBP-SB protein binds to DNA as a monomer and is not able to form dimers *in vitro*. This can be due to improper folding of the transposase caused by the large MBP tag. The functionality of the protein must be disrupted as in the *in vivo* transposition assay no activity of the fusion transposase was proved. Transfection of the MBP-SB with neo-resistant transposon construct did not lead to a raise in G-418 resistant colony number in comparison to the negative control performed with catalytically inactive transposase. Lack of the *in vivo* activity of the MBP-SB fusion transposase can be the explanation for the failure of the *in vitro* trials.

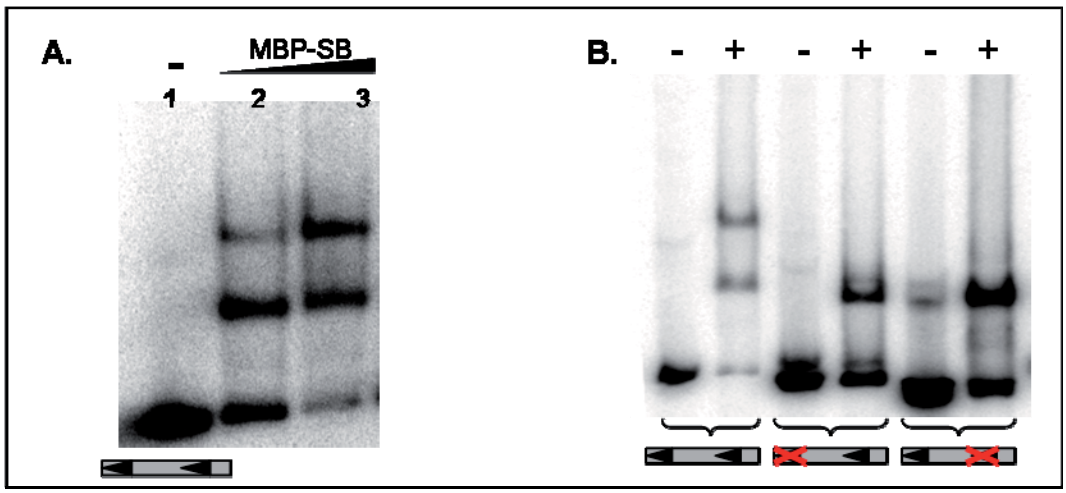

Abb. 12: MBP is able to bind DNA but not to form dimers. A. demonstration of the DNA binding ability of the MBP-SB in EMSA (lane 2 and 3); **B.** MBP-SB forms single complexes with IRs lacking one binding site; - and + stand for without and with protein, respectively.

In contrary, the His-tagged *SB* had activity *in vivo*, though, very low. The purified protein, on the other hand, was not stable as I observed degradation products (data not shown). Additionally, in binding assays only bands migrating as purified His-tagged DNA-binding domain of the *SB* used as a control for the assay were present indicating formation of complexes with shorter than full-length proteins (data not shown). As it seemed that the size of the tag may distort the proper folding of the transposase and this way inactivate it and fusion with short tag retained some activity, another small tag was fused to the transposase. STREP is an eight-residue peptide sequence that exhibits intrinsic affinity toward streptavidin. The expressed STREP-SB fusion protein (Abb. 13A, lane 2) goes into inclusion bodies, as after sonication no *SB* was detected in the supernatant fraction but a strong band corresponding to the *SB* transposase was present in the pellet fraction (Abb. 13A, lane 3 and 4, respectively). The complete insolubility of the fusion protein was the obstacle in obtaining functional *SB* transposase.
The chemical properties of the protein dramatically changed when between the STREP-tag and the *SB* a JAZZ linker was introduced. JAZZ is an artificial zinc finger protein binding to a 9-bp sequence in the promoter region of the human utrophin gene [Corbi N, 2000]. The idea of fusing the JAZZ linker to the *SB* transposase comes from observations of my colleague. Her JAZZ-SB fusion was able to promote the transposition reaction *in vivo*. When the STREP was added for purification purposes the STREP-JAZZ-SB construct retained moderate activity in HeLa cells, estimated from three independent transposition assay experiments for approximately 20% of the wild-type transposase (data not shown). Therefore, the purified STREP-JAZZ-SB was subjected to the *in vitro* transposition assay (shown in Abb. 11).

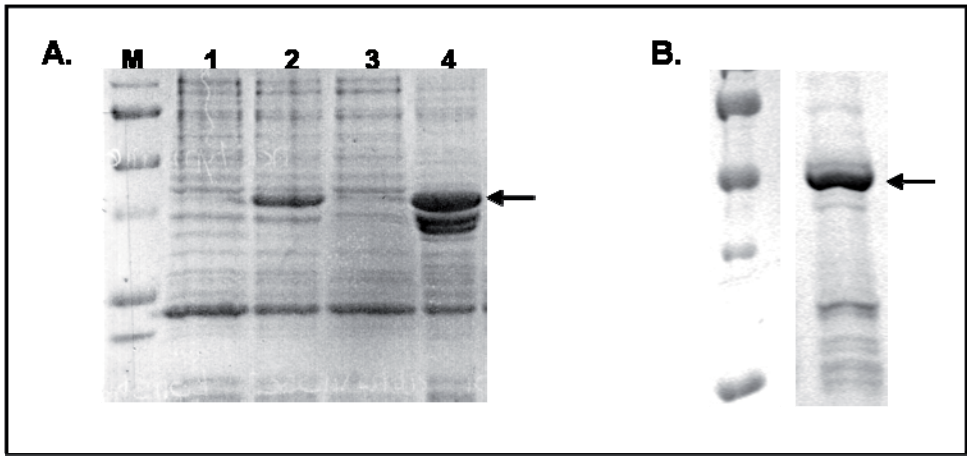

Abb. 13: Purification of the SB fusion proteins. A. Coomassie staining of the purifications steps of the STREP-SB. On the gel fractions of the uninduced (1) and induced (2) *E.coli*, supernatant (3) and pellet (4) after sonication. M stands for the marker. **B.** Purified STREP-JAZZ-SB on the polyacrylamide gel stained with Coomassie.

Conditions for the *SB in vitro* transposition were chosen based on published data on the *in vitro* systems of the related Tc1 and *mariner* transposons [Vos, 1996], [Lampe, 1996] and knowledge collected in our laboratory about *Sleeping Beauty* transposase and properties of the system. The basic buffer for the reaction was composed of 25 mM HEPES and 100 mM KCl supplied with 5 mM $MgCl_2$ and 1 mM DTT. Additionally, reactions were conducted in enriched buffers containing 10 % glycerol, 10 % DMSO or 0.02 % NP40. The target and donor DNA plasmid were supplied in amounts of 2 µg and 200 ng, respectively, and the purified STREP-JAZZ-SB at low concentration to avoid the overproduction inhibition effect. Purified HMGB1, shown to be the host factor of the transposition was also added to the reaction. Moreover, as no colonies growing on Amp+Zeo+Strep and sucrose selection were recovered, reactions with different protein-DNA ratios were performed to find optimal conditions for the transposition to take place. However, no transposition events were detected in the *in vitro* assays with the conditions applied.

Approaches to obtain an active SB-fusion protein and properties and activities of the created fusions are summarized in Tab. 1. Unfortunately, none of the fusion proteins was successful and trials to establish the *in vitro* transposition assay failed.

	His-SB	MBP-SB	STREP-SB	STREP-JAZZ-SB	HA-SB
Solubility	Very low	√	-	good	√
DNA binding	Partially	√	nd	good	?
In vivo activity	15%	-	nd	20%	50%
In vitro activity	nd	nd	nd	no activity	?

Tab. 1: Properties and activities of the tagged SB transposases. In blue the activities and properties tested within the project. * only binding of the N-terminal domain to DNA could be detected.

6 DISCUSSION

Transposons are mobile pieces of DNA widespread throughout all living organisms. They can be grouped in families based on their structure and mechanism of transposition (Abb. 1). The transposition process is believed to be a multistep reaction with several protein-DNA and protein-protein interactions that take place in a sequential manner. The reaction can be regulated at different steps by transposon- and host-mediated mechanisms. *Sleeping Beauty* is a fish transposon reconstructed from molecular fossils [Ivics, 1997] to restore a vertebrate transposable element for genetic applications. It is a member of Tc1/*mariner* family with a complex IR/DR structure of the transposon DNA. The IR/DR structure refers to inverted repeats (IR) each containing two direct repeats (DR) serving as binding sites for the transposase (Abb. 2B). The inverted repeat upstream of the gene is commonly called left IR and the inverted repeat downstream of the gene, right IR. The left IR contains additionally between the DRs an enhancer site. The binding sites are very similar but not identical, showing 3 nucleotides differences in sequence and 2 nucleotides differences in length. The complex IR/DR structure of *SB,* with its sequence, combination of the inter-DR spacing and geometry of the DRs, is very sensitive to changes as exchanging the inner DRs for the outer DRs or removing the sequence between the DRs strongly decreased the transposition efficiency [Cui, 2002], [Izsvak, 2002]. These observations indicate that the IR/DR structure is of high importance for the transposition reaction. Moreover, the results of experiments aiming to find sequences bound stronger by the transposase showed that increasing the strength of the transposase binding to the transposon DNA does not boost the overall transposition rate. Quite the contrary, enhancing one step of the reaction may negatively influence the following steps leading to decreased transposition efficiency. One explanation can be that the assembly of the four binding sites and transposase molecules is not a spontaneous, but a precisely arranged process. Altering the DNA binding affinities of the binding sites could perturb the strict order of the assembly process.

In order to understand the function of the complex IRs of the *SB*, I compared the system to a transposon with simple IRs. The IR/DR structure of the *Sleeping Beauty* transposon is not present in the *mariner*-like elements, a subfamily of the Tc1/*mariner* transposons. *mariners* have short IRs each containing one binding site for the transposase. Both systems were subjected to *in vivo* transposition assay using complete transposon sequences and constructs lacking the right IRs. Dimer of the *Himar1* transposase was previously reported to show catalytic activities on one transposon end [Lipkow, 2004]. Similarly, many transposition events were recovered with single-end *Hsmar1* transposon (Abb. 5A and B). *Sleeping Beauty* was also capable of single-end transposition but to much lower extent. During such transposition, the borders of the transposon are not precisely defined. While in *SB* transposition, the cleavage and transposition of the left IR is normal, the liberation of the other end of the transposon does not seem to be transposase-mediated. In principle, the second DNA break can occur inside of the transposon

or, somewhere in the genome, contributing this way to genomic instability. Lower number of imprecise transposition events is a result of better regulation of the *SB* transposition that might be imposed by the IR/DR stucture. The complexity of the inverted repeats together with the bipartite (PAIRED) DNA-binding domain of *SB* might offer means for more sophisticated regulation of the transposition process, thereby protecting the genome from aberrant transposition events.

6.1 Dissection of the paired-end complex formation

I performed experiments to dissect the transposition process and to gain insight into single steps of the reaction. When the subdomains were separated and analyzed independently from each other, I could learn not only about their activities but the results allowed me to draw a model how the transposition complex is formed.
PAI subdomain of the transposase specifically recognizes the binding substrate. The specificity of self-identification is high as *SB* is able to distinguish its own sequence from a related zebrafish binding site, differing only by 5 out of 17 bp in the PAI recognition motif [Ivics, 1997]. Incubating purified PAI and RED subdomains with oligonucleotides corresponding to each binding site and resolving in electrophoretic mobility shift assay (EMSA) revealed that the PAI binds stronger the inner binding site (Abb. 6A, lanes 3 and 6). Interestingly, the RED subdomain binds only the inner binding site. No binding of the RED to the outer binding site was detected (Abb. 6A, lanes 2 and 5). Results of the CAST-ing experiment suggested that only the PAI subdomain is involved in the primary substarte binding, as only the PAI recognition motif of the DRs was present in the CAST-ing selected oligonucleotides while the RED recognition motif was random. However, as the RED showed no recognition of the outer binding site it could suggest that the function of the RED subdomain in the primary binding would be to enhance the binding towards the inner binding site. It seems that the RED might be responsible for determining the primary binding target, which is the inner binding site. Starting the complex formation from the inner binding site, distantly from the cleavage site, can be a manner to prevent enzymatic activities such as nicking and cleavage before the paired-end-complex (PEC) is formed. The RED binding is very specific and restricted to the inner binding site, as introducing two nucleotides that are present in the outer binding site mimicking the longer binding site but leaving the rest of the sequence unmodified resulted in disruption of the binding (Abb. 6B, lanes 2 and 4). Additionally, I concluded from this EMSA experiment that the RED subdomain formed a dimer on the binding site. As I never observed a band corresponding with size to a RED monomer bound to the DNA it is possible that the protein subdomain binds in a dimeric form. For *mariner* transposases, dimers appear to be the form binding to the inverted repeats [Augé-Gouillou, 2005]. On the other hand, Butler and colleagues argue that *mariners* might bind as monomers and then quickly recruit another monomer to form a very

stable dimer [Butler, 2006]. Their statement is supported by the observation of the *Himar1* transposase mutant lacking the protein-protein interaction activity that was still able to bind the DNA. Similarly, it is highly probable that the monomer of RED subdomain binds to the direct repeat. Then, the binding initiates a conformational change what releases protein-protein interaction activity of the RED leading to dimer formation. It can not be excluded that the active form of the RED subdomain able to interact with DNA is a dimer; however, binding of the monomer followed by dimerization can be supported by the fact that the full-length DNA-binding domain was reported to be in a monomeric state in solution and tetramerize in the presence of DNA [Izsvak, 2002].
Interestingly, even though the outer direct repeat was not recognized by the RED subdomain in binding assays, its sequence can not be freely changed suggesting its importance in transposition. Footprinting data have shown that RED subdomain binds the 5'-part of the DRs and PAI the 3'-part [Izsvak, 2002]. In case of the outer binding sites RED contacts regions of the inverted repeats close to the cleavage sites. Substituting the RED recognition motif of the outer binding site with the sequence of one of the oligonucleotides selected in the CAST-ing experiment showing no significant similarities to the wild-type binding site resulted in decreased transposition efficiency (Abb. 7). As the transposition reaction consists of successive protein-protein and protein-DNA interactions, the RED-outer DR interaction may have a function at later step of the process, e.g. in cleavage. *mariner*-like element, *Himar1* transposase containes only one HTH motif placed in the C terminus of the DNA-binding domain. Mutagenesis of this region of the *Himar1* DNA-binding domain lead to impairment of cleavage and strand transfer [Butler, 2006]. If there is analogy of the structural and functional organization of the DNA-binding domain of the *mariner*-like elements and *Sleeping Beauty*, the cleavage and insertion *Himar1* mutants may support the possible role of the RED subdomain in the cleavage upon interaction with the 5'-part of the outer DRs. Thus, no interaction of the RED with the outer binding site at the early steps of the transposition process additionally counteracts cleavage before the proper PEC formation.
Since dimers of the *mariner* transposase formed on a single transposon end show catalytic activities [Lipkow, 2004] I checked whether the dimer of the SB transposase bound to the inner binding site or bringing two inner binding sites together is able to cleave the DNA at the inner binding sites. In order to answer this question, a truncated transposon construct was subjected to *in vivo* excision assay. The left inverted repeat was unmodified as removal of one DR from the left IR results in abolishment of the transposition (Abb. 4, construct 1). In the right IR only the inner DR was present, followed by the TA to mimic the cleavage site. No excision product was detected in the PCR based *in vivo* excision assay (Abb. 10C, construct T/in). These results suggest existence of a regulatory mechanism inhibiting cleavage at the inner binding site, within the transposon sequence. Thus, unlike the *mariners* in *SB* system, the catalytic steps are prevented before the PEC formation.

6.2 Influence of host factor HMGB1 on PEC formation

Sleeping Beauty is active in many cell lines of different organisms suggesting no requirements for species specific host factors [Izsvák, 2000]. Nevertheless, an affect of host factors cannot be excluded. Indeed, four host factors have been identified that influence the transposition of *SB* and other transposable elements [Zayed, 2003], [Izsvak, 2004], [Walisko, 2006]. Many DNA recombinational processes require beside the recombinase additional proteins altering the DNA conformation. The complex assembly of the Mu transpososome does not proceed in absence of HU protein bringing the binding sites of the Mu transposase together [Lavoie, 1993]. HMGB1 is a eukaryotic protein bending the DNA [Thomas, 2001] previously reported to enhance the *SB* binding towards the inner binding site and to conduce the bending of the DNA bringing the transposon ends in proximity [Zayed, 2003]. Even though, in my experiments oligonucleotides corresponding to the binding sites and not the full-length transposon DNA were used, the positive influence of the HMGB1 was observed. Presence of the HMGB1 lead to enhanced formation of the complex containing dimer of the RED subdomain binding two inner DRs (Abb. 9A, compare lanes 2-4 and 5-7). The mechanism of the enhancement is, however, not clear. The bending protein is not stably incorporated into the complex as no additional shifted band appeared upon incubation with HMGB1. Similar observations were collected with the full-length *SB* transposase (MBP-SB, [Zayed, 2003]; my data not shown). HMGB1 interaction with the RED subdomain was shown in the chemical cross-linking experiment (Abb. 9B lane 3).

6.3 Model of PEC formation

I propose that the IR/DR structure with its four transposase binding sites of diverse affinities for the transposase at single steps of the reaction evolved to arrange the formation of the synaptic complex, which is the prerequisite for the subsequent catalytic steps of transposition to take place.
Based on my results a model of the paired-end complex formation was developed (Abb. 14). In the overall transposition process of SB, the left IR might be preferred over the right one, since left IR mutant/deleted transposons are completely inactive for transposition, while transposition can proceed even in the lack of complete right IR (Abb. 4, constructs 2 and 4). Interestingly, the left IR can be duplicated without major disturbance of the transposition reaction [Izsvak, 2002]. The affinity of the transposase is higher towards the inner binding site due to the RED subdomain distinguishing between the inner and outer binding site. This phenomenon is additionally enhanced by the host factor HMGB1. Thus, the complex formation starts at the inner DR located distantly from the cleavage site. The dimer of the protein formed on the left inner binding site is able to capture the second transposon end. Incorporation of the inner binding site from the right in-

verted repeat results in PEC formation. Importantly, the inner binding site is resistant for cleavage even if the canonical cleavage site (TA|CAGT) is artificially engineered nearby (Abb. 10). Accordingly, replacing the outer binding site by the inner binding site at the end of the transposon, even if the cleavage site is available, decreases transposition by 80% [Cui, 2002]. Thus, starting the PEC formation with the inner binding sites decreases the probability of cleavage before the proper complex fromation.
The order of binding and inclusion of the other two binding sites into the PEC is not known in detail, but the incorporation of the outer DRs takes place after the inner DRs are brought together. The protein-protein and the DNA-interaction interfaces of PAI and RED are proposed to be utilized in an alternative manner at different steps of the assembly process. For example, the homo-dimerization of RED but not PAI could be detected at the early steps of PEC formation. The protein interaction interface of PAI is predicted to play a role at a later step of the complex assembly, perhaps in the process of positioning the IRs for cleavage. The IR/DR structure of *SB* seems to enforce PEC formation, and decrease the frequency of imprecise transposition events.
For some transposons, the complex formation has been studied very deeply and there is evidence that it is a multistep complex reaction. The paired-end-complex formation, where two transposon ends are brought in close proximity, is a prerequisite for the catalytic activities of the bacterial Tn5 and Tn10 transposons as the transposases act in *trans* [Naumann, 2000]. It means that the transposase bound to one transposon end process the cleavage of the second transposon end Activity on a single end of the transposon is inhibited because the transposase is unable to dimerize in the absence of the second transposon end [Naumann, 2000]. Thus, if proper complex formation fails, no excision would take place. For all of classical bacterial transposons characterized to date, the catalytic steps of the reaction are tightly coupled to the synapsis of the transposon ends (PEC) [Savilahti, 1996]. Therefore, PEC formation serves an important regulatory checkpoint of the transposition process.
In contrary, *mariner*-like element *Himar1* was reported to form active complexes on one transposon end *in vitro* [Lipkow, 2004]. Thus, in *mariner*s systems, PEC formation does not serve as an absolute checkpoint control to enforce the synapsis of the two ends of the transposon prior catalysis. *Sleeping Beauty* was also able to transpose with only one inverted repeat; however, with much lower efficiency than *Hsmar1* (Abb. 5). The organization of the IRs, short with single binding sites for *mariners* and long IRs each containing two DRs of *Sleeping Beauty*, implies that different pathways are used by *mariner*- and Tc1-like elements to regulate transposition initiation and thus leading to the formation of differently organized synaptic complexes. My results indicate that the IR/DR structure of the *SB* transposon introduces higher level of regulation of the transposition process. Four binding sites of different affinities and two DNA-binding subdomains serve additional regulation points for the ordered assembly process. “Ordered assembly” as a regulatory mechanism was already described in the context of V(D)J recombina-

tion or bacteriophage *Mu* transposition. Such regulation is able to filter out undesired or aberrant reaction intermediates.

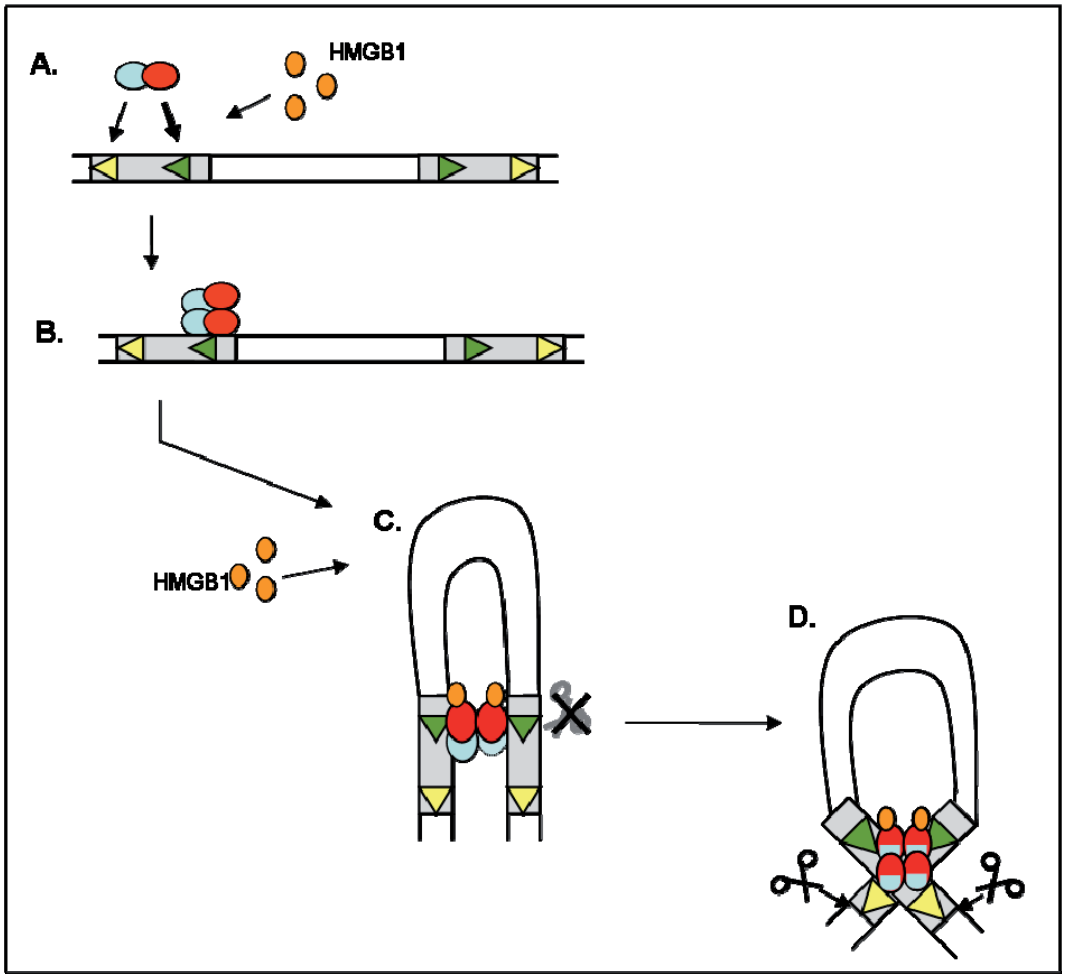

Abb. 14: Model of the paired-end-complex (PEC) formation. The blue and red ovals represent the PAI and the RED subdomain of the *SB* DNA-binding domain, respectively. **A.** The DNA-binding domain of the transposase is able to bind both binding sites but the affinity towards the inner binding site is enhanced. The formation of the complex starts at the inner DR. This step is enhanced by the HMGB1 (EMSA with the full-length *SB*; [**Zayed, 2003**]). **B.** The next step is the dimerization of the SB transposase performed by the RED subdomain. **C.** The RED is then able to catch the second transposon end, which is the second inner binding site. This step is enhanced by the HMGB1 (here shown as orange ovals). **D.** Incorporation of two transposase monomers interacting with the outer binding sites accomplishes the PEC formation. The scissors imply that the complex is ready to undergo catalytic activities.

6.4 *In vitro* transposition assay

The *in vitro* transposition assays have been developed for the mariner *Mos1* [Tosi, 2000] and Tc1 element [Vos, 1996], both having short inverted repeats. These transposons are able to mediate the reaction in simple conditions as long as the donor and target DNA, the transposase and the magnesium divalent ions for protein's catalytic activities are supplied. The results from the Tc1 experiments indicate possible stimulatory role of the host factors on the transposition reaction as the assay with the worm extract yielded significantly higher number of transposition events than the assay with Tc1 transposase purified from the E.coli. Nevertheless, transposase is the only protein required and the transposition can take place in suboptimal conditions. These conditions were applied for the *Sleeping Beauty in vitro* transposition assay; however, with no positive result. The assay was modified based on our knowledge about the *SB* system collected until now to apply the probable preferred conditions. Those include low amounts of the transposase as *SB* undergoes overproduction inhibition, meaning that high concentration of the protein leads to decrease of the transposition efficiency, and addition of the host factor HMGB1. Although different protein-DNA ratios were used no transposition event was detected. It can be due to the fusion proteins used in the experiments. *Sleeping Beauty* transposase does not tolerate tags on the C-terminus and the tags on the N-terminus significantly decrease the activity of the protein. Thus, it can not be excluded that the purified proteins did not exhibit all the required activities. However, the failure of the establishment of the *in vitro* transposition assay may result from the sophisticated regulation of the multistep reaction that will not proceed to the next step if the conditions are not fulfilled.

REFERENCES

Aidinis, 1999 Aidinis V, Bonaldi T, Beltrame M, Santagata S, Bianchi ME, Spanopoulou E: The RAG1 homeodomain recruits HMG1 and HMG2 to facilitate recombination signal sequence binding and to enhance the intrinsic DNA-bending activity of RAG1-RAG2. Mol Cell Biol. 1999, *19(10)*, S.6532-42,

Augé-Gouillou, 2005 Augé-Gouillou C, Brillet B, Germon S, Hamelin MH, Bigot Y: Mariner Mos1 transposase dimerizes prior to ITR binding. J Mol Biol. 2005, *351(1)*, S.117-30,

Bannert, 2004 Bannert N, Kurth R: Retroelements and the human genome: new perspectives on an old relation. Proc Natl Acad Sci U S A. 2004, *101 Suppl 2*, S.14572-9,

Berg, 1982 Berg DE, Johnsrud L, McDivitt L, Ramabhadran R, Hirschel BJ: Inverted repeats of Tn5 are transposable elements. Proc Natl Acad Sci U S A. 1982, *79(8)*, S.2632-5,

Beuzón, 1999 Beuzón CR, Marqués S, Casadesús J: Repression of IS200 transposase synthesis by RNA secondary structures. Nucleic Acids Res. 1999, *27(18)*, S.3690-5,

Buschang, 2001 Buschang PH, Throckmorton GS, Travers KH, Hayasaki H: Incisor and mandibular condylar movements of young adult females during maximum protrusion and lateratrusion of the jaw. Arch Oral Biol. 2001, *46(1)*, S.39-48,

Bustin, 1999 Bustin M: Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. Mol Cell Biol. 1999, *19(8)*, S.5237-46,

Butler, 2006 Butler MG, Chakraborty SA, Lampe DJ: The N-terminus of Himar1 mariner transposase mediates multiple activities during transposition. Genetica. 2006, *127(1-3)*, S.351-66,

Chaconas, 1988 Chaconas G, Surette MG: Mechanism of Mu DNA transposition. Bioessays. 1988, *9(6)*, S.205-8,

Chaconas, 1996 Chaconas G, Lavoie BD, Watson MA: DNA transposition: jumping gene machine, some assembly required. Curr Biol. 1996, *6(7)*, S.817-20,

Chandler, 1993 Chandler M, Fayet O: Translational frameshifting in the control of transposition in bacteria. Mol Microbiol. 1993, *7(4)*, S.497-503,

Colloms, 1994 Colloms SD, van Luenen HG, Plasterk RH: DNA binding activities of the Caenorhabditis elegans Tc3 transposase. Nucleic Acids Res. 1994, *22(25)*, S.5548-54,

Corbi N, 2000 Corbi N, Libri V, Fanciulli M, Tinsley JM, Davies KE, Passananti C: The artificial zinc finger coding gene 'Jazz' binds the utrophin promoter and activates transcription. Gene Ther. 2000, *7(12)*, S.1076-83,

Cui, 2002 Cui Z, Geurts AM, Liu G, Kaufman CD, Hackett PB: Structure-function analysis of the inverted terminal repeats of the sleeping beauty transposon. J Mol Biol. 2002, *318(5)*, S.1221-35,

Czerny, 1993 Czerny T, Schaffner G, Busslinger M: DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. Genes Dev. 1993, *7(10)*, S.2048-61,

Davies, 2000 Davies DR, Goryshin IY, Reznikoff WS, Rayment I: Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. Science. 2000, *289(5476)*, S.77-85,

Devine, 1994 Devine SE, Boeke JD: Efficient integration of artificial transposons into plasmid targets in vitro: a useful tool for DNA mapping, sequencing and genetic analysis. Nucleic Acids Res. 1994, *22(18)*, S.3765-72,

Difilippantonio, 1996 Difilippantonio MJ, McMahan CJ, Eastman QM, Spanopoulou E, Schatz DG: RAG1 mediates signal sequence recognition and recruitment of RAG2 in V(D)J recombination. Cell. 1996, *87(2)*, S.253-62,

Doak, 1994 Doak TG, Doerder FP, Jahn CL, Herrick G: A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. Proc Natl Acad Sci U S A. 1994, *91(3)*, S.942-6,

Emmons, 1983 Emmons SW, Yesner L, Ruan KS, Katzenberg D: Evidence for a transposon in Caenorhabditis elegans. Cell. 1983, *32(1)*, S.55-65,

Essner, 2005 Essner JJ, McIvor RS, Hackett PB: Awakening gene therapy with Sleeping Beauty transposons. Curr Opin Pharmacol. 2005, *5(5)*, S.513-9,

Fedoroff, 1983 Fedoroff N, Wessler S, Shure M.: Isolation of the transposable maize controlling elements Ac and Ds. Cell . 1983, *35(1)*, S.235-42,

Feng, 1994 Feng JA, Johnson RC, Dickerson RE.: Hin recombinase bound to DNA: the origin of specificity in major and minor groove interactions. Science. 1994, *263(5145)*, S.348-55,

Feng, 2007 Feng X, Colloms SD.: In vitro transposition of ISY100, a bacterial insertion sequence belonging to the Tc1/mariner family. Mol Microbiol. 2007, *65(6)*, S.1432-43,

Fischer, 1999 Fischer SE, van Luenen HG, Plasterk RH: Cis requirements for transposition of Tc1-like transposons in C. elegans. Mol Gen Genet. 1999, *262(2)*, S.268-74,

Fischer, 2001 Fischer SE, Wienholds E, Plasterk RH: Regulated transposition of a fish transposon in the mouse germ line. Proc Natl Acad Sci U S A. 2001, *98(12)*, S.6759-64,

Foster, 1981 Foster TJ, Davis MA, Roberts DE, Takeshita K, Kleckner N: Genetic organization of transposon Tn10. Cell. 1981, *23(1)*, S.201-13,

Gehring, 1994 Gehring WJ, Qian YQ, Billeter M, Furukubo-Tokunaga K, Schier AF, Resendez-Perez D, Affolter M, Otting G, Wüthrich K: Homeodomain-DNA recognition. Cell. 1994, *78(2)*, S.211-23,

Geurts, 2003 Geurts AM, Yang Y, Clark KJ, Liu G, Cui Z, Dupuy AJ, Bell JB, Largaespada DA, Hackett PB: Gene transfer into genomes of human cells by the sleeping beauty transposon system. Mol Ther. 2003, *8(1)*, S.108-17,

Goins, 2008 Goins WF, Krisky DM, Wechuck JB, Huang S, Glorioso JC: Construction and production of recombinant herpes simplex virus vectors. Methods Mol Biol. 2008, *433*, S.97-113,

Goryshin, 1998 Goryshin IY, Reznikoff WS: Tn5 in vitro transposition. J Biol Chem. 1998, *273(13)*, S.7367-74,

Gunzburg, 1996 Gunzburg WH, Salmons B: Development of retroviral vectors as safe, targeted gene delivery systems. J Mol Med. 1996, *74(4)*, S.171-82,

Hacein-Bey-Abina, 2003 Hacein-Bey-Abina, S., C. von Kalle, M. Schmidt, M. P. McCormack, N. Wulffraat, P. Leboulch, A. Lim, C. S. Osborne, R. Pawliuk, E. Morillon, R. Sorensen, A. Forster, P. Fraser, J. I. Cohen, G. de Saint Basile, I. Alexander, U. Wintergerst, T. Frebourg, A. Aurias, D. Stoppa-Lyonnet, S. Romana, I. Radford-Weiss, F. Gross, F. Valensi, E. Delabesse, E. MacIntyre, F. Sigaux, J. Soulier, L. E. Leiva, M. Wissler, C. Prinz, T. H. Rabbitts, F. Le Deist, A. Fischer, and M. Cavazzana-Calvo: LM02-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science . 2003, *302*, S.415-419,

He, 2004 He CX, Shi D, Wu WJ, Ding YF, Feng DM, Lu B, Chen HM, Yao JH, Shen Q, Lu DR, Xue JL: Insulin expression in livers of diabetic mice mediated by hydrodynamics-based administration. World J Gastroenterol. 2004, *10(4)*, S.567-72,

Henikoff, 1992 Henikoff S: Detection of Caenorhabditis transposon homologs in diverse organisms. New Biol. 1992, *4(4)*, S.382-8,

Houwing, 2007 Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB, Plasterk RH, Hannon GJ, Draper BW, Ketting RF: A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell. 2007, *129(1)*, S.69-82,

Ivics, 1996 Ivics Z, Izsvak Z, Minter A, Hackett PB: Identification of functional domains and evolution of Tc1-like transposable elements. Proc Natl Acad Sci U S A. 1996, *93(10)*, S.5008-13,

Ivics, 1997 Ivics Z, Hackett PB, Plasterk RH, Izsvak Z: Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell. 1997, *91(4)*, S.501-10,

Ivics, 1999 Ivics Z, Izsvak Z, Hackett PB: Genetic applications of transposons and other repetitive elements in zebrafish. Methods Cell Biol. 1999, *60*, S.99-131,

Izsvák, 2000 Izsvák Z, Ivics Z, Plasterk RH: Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. J Mol Biol. 2000, *302(1)*, S.93-102,

Izsvak, 2002 Izsvak Z, Khare D, Behlke J, Heinemann U, Plasterk RH, Ivics Z: Involvement of a bifunctional, paired-like DNA-binding domain and a transpositional enhancer in Sleeping Beauty transposition. J Biol Chem. 2002, *277(37)*, S.34581-8,

Izsvak, 2004 Izsvak Z, Stuwe EE, Fiedler D, Katzer A, Jeggo PA, Ivics Z: Healing the wounds inflicted by sleeping beauty transposition by double-strand break repair in mammalian somatic cells. Mol Cell. 2004, *13(2)*, S.279-90,

Jacobson, 1985 Jacobson JW, Hartl DL: Coupled instability of two X-linked genes in Drosophila mauritiana: germinal and somatic mutability. Genetics. 1985, *111(1)*, S.57-65,

Jun, 1996 Jun S, Desplan C: Cooperative interactions between paired domain and homeodomain. Development. 1996, *122(9)*, S.2639-50,

Junop, 1997 Junop MS, Haniford DB: Factors responsible for target site selection in Tn10 transposition: a role for the DDE motif in target DNA capture. EMBO J. 1997, *16(10)*, S.2646-55,

Kaufman, 1992 Kaufman PD, Rio DC: P element transposition in vitro proceeds by a cut-and-paste mechanism and uses GTP as a cofactor. Cell. 1992, *69(1)*, S.27-39,

Kennedy, 1997 Kennedy PG: Potential use of herpes simplex virus (HSV) vectors for gene therapy of neurological disorders. Brain. 1997, *120 (Pt 7)*, S.1245-59,

Kleckner, 1996 Kleckner N, Chalmers RM, Kwon D, Sakai J, Bolland S: Tn10 and IS10 transposition and chromosome rearrangements: mechanism and regulation in vivo and in vitro. Curr Top Microbiol Immunol. 1996, *204*, S.49-82,

Krause, 1986 Krause HM, Higgins NP: Positive and negative regulation of the Mu operator by Mu repressor and Escherichia coli integration host factor. J Biol Chem. 1986, *261(8)*, S.3744-52,

Kunze, 1993 Kunze R, Behrens U, Courage-Franzkowiak U, Feldmar S, Kuhn S, Lutticke R: Dominant transposition-deficient mutants of maize Activator (Ac) transposase. Proc Natl Acad Sci U S A. 1993, *90(15)*, S.7094-8,

Kunze, 1996 Kunze R: The maize transposable element activator (Ac). Curr Top Microbiol Immunol. 1996, *204*, S.161-94,

Lampe, 1996 Lampe DJ, Churchill ME, Robertson HM: A purified mariner transposase is sufficient to mediate transposition in vitro. EMBO J. 1996, *15(19)*, S.5470-9,

Lampe, 1999 Lampe DJ, Akerley BJ, Rubin EJ, Mekalanos JJ, Robertson HM: Hyperactive transposase mutants of the Himar1 mariner transposon. Proc Natl Acad Sci U S A. 1999, *96(20)*, S.11428-33,

Lander, 2001 Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL,

Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ; International Human Genome Sequencing Consortium: Initial sequencing and analysis of the human genome. Nature. 2001, *409(6822)*, S.860-921,

Lavoie, 1993 Lavoie BD, Chaconas G: Site-specific HU binding in the Mu transpososome: conversion of a sequence-independent DNA-binding protein into a chemical nuclease. Genes Dev. 1993, *7(12B)*, S.2510-9,

Leung, 1989 Leung PC, Teplow DB, Harshey RM: Interaction of distinct domains in Mu transposase with Mu DNA ends and an internal transpositional enhancer. Nature. 1989, *338(6217)*, S.656-8,

Lipkow, 2004 Lipkow K, Buisine N, Lampe DJ, Chalmers R: Early intermediates of mariner transposition: catalysis without synapsis of the transposon ends suggests a novel architecture of the synaptic complex. Mol Cell Biol. 2004, *24(18)*, S.8301-11,

Lohe, 1996 Lohe AR, Hartl DL: Autoregulation of mariner transposase activity by overproduction and dominant-negative complementation. Mol Biol Evol. 1996, *13(4)*, S.549-55,

Lohe, 1997 Lohe AR, De Aguiar D, Hartl DL: Mutations in the mariner transposase: the D,D(35)E consensus sequence is nonfunctional. Proc Natl Acad Sci U S A. 1997, *94(4)*, S.1293-7,

Luo, 1998 Luo G, Ivics Z, Izsvak Z, Bradley A: Chromosomal transposition of a Tc1/mariner-like element in mouse embryonic stem cells. Proc Natl Acad Sci U S A. 1998, *95(18)*, S.10769-73,

Luthi, 1990 Luthi K, Moser M, Ryser J, Weber H: Evidence for a role of translational frameshifting in the expression of transposition activity of the bacterial insertion element IS1. Gene. 1990, *88(1)*, S.15-20,

Masson, 1991 Masson P, Strem M, Fedoroff N: The tnpA and tnpD gene products of the Spm element are required for transposition in tobacco. Plant Cell. 1991, *3(1)*, S.73-85,

Mátés, 2009 Mátés L, Chuah .M.K.L., Belay E, Jerchow B, Manoj N, Acosta-Sanchez A, Grzela D, Schmitt A, Becker K, Matrai J, Ma L, Samara-Kuko E, Gysemans C, Pryputniewicz D, Miskey C, Fletcher B, VandenDriessche T, Ivics Z, Izsvák ZMolecular Evolution of a Novel Hyperactive Sleeping Beauty Transposase Enables Robust Stable Gene Transfer in VertebratesNat Genet200*41(6)*753-61

Mátés, 2007 Mátés L, Izsvák Z, Ivics Z: Technology transfer from worms and flies to vertebrates: transposition-based genome manipulations and their future perspectives. Genome Biol. 2007, *8 Suppl 1*, S.S1,

McClintock, 1950 McClintock B: The origin and behavior of mutable loci in maize. Proc Natl Acad Sci U S A. 1950, *36(6)*, S.344-55,

Miskey, 2003 Miskey C, Izsvák Z, Plasterk RH, Ivics Z: The Frog Prince: a reconstructed transposon from Rana pipiens with high transpositional activity in vertebrate cells. Nucleic Acids Res. 2003, *31(23)*, S.6873-81,

Miskey, 2007 Miskey C, Papp B, Mátés L, Sinzelle L, Keller H, Izsvák Z, Ivics Z: The ancient mariner sails again: transposition of the human Hsmar1 element by a reconstructed transposase and activities of the SETMAR protein on transposon ends. Mol Cell Biol. 2007, *27(12)*, S.4589-600,

Montini, 2002 Montini E, Held PK, Noll M, Morcinek N, Al-Dhalimy M, Finegold M, Yant SR, Kay MA, Grompe M: In vivo correction of murine tyrosinemia type I by DNA-mediated transposition. Mol Ther. 2002, *6(6)*, S.759-69,

Naumann, 2000 Naumann TA, Reznikoff WS: Trans catalysis in Tn5 transposition. Proc Natl Acad Sci U S A. 2000, *97(16)*, S.8944-9,

Naumann, 2002 Naumann TA, Reznikoff WS: Tn5 transposase active site mutants. J Biol Chem. 2002, *277(20)*, S.17623-9,

Nekrutenko, 2001 Nekrutenko A, Li WH: Transposable elements are found in a large number of human protein-coding genes. Trends Genet. 2001, *17(11)*, S.619-21,

Nolan, 2005 Nolan T, Braccini L, Azzalin G, De Toni A, Macino G, Cogoni C: The post-transcriptional gene silencing machinery functions independently of DNA methylation to repress a LINE1-like retrotransposon in Neurospora crassa. Nucleic Acids Res. 2005, *33(5)*, S.1564-73,

Ortiz-Urda, 2003 Ortiz-Urda S, Lin Q, Yant SR, Keene D, Kay MA, Khavari PA: Sustainable correction of junctional epidermolysis bullosa via transposon-mediated nonviral gene transfer. Gene Ther. 2003, *10(13)*, S.1099-104,

Park, 2008 Park K, Kim WJ, Cho YH, Lee YI, Lee H, Jeong S, Cho ES, Chang SI, Moon SK, Kang BS, Kim YJ, Cho SH: Cancer gene therapy using adeno-associated virus vectors. Front Biosci. 2008, *13*, S.2653-9,

Peukert, 1997 Peukert K, Staller P, Schneider A, Carmichael G, Hänel F, Eilers M: An alternative pathway for gene regulation by Myc. EMBO J. 1997, *16(18)*, S.5672-86,

Plasterk, 1996 Plasterk RH: The Tc1/mariner transposon family. Curr Top Microbiol Immunol. 1996, *204*, S.125-43,

Plasterk, 1999 Plasterk RH, Izsvák Z, Ivics Z: Resident aliens: the Tc1/mariner superfamily of transposable elements. Trends Genet. 1999, *15(8)*, S.326-32,

Pledger, 2005 Pledger DW, Coates CJ: Mutant Mos1 mariner transposons are hyperactive in Aedes aegypti. Insect Biochem Mol Biol. 2005, *35(10)*, S.1199-207,

Polard, 1991 Polard P, Prère MF, Chandler M, Fayet O: Programmed translational frameshifting and initiation at an AUU codon in gene expression of bacterial insertion sequence IS911. J Mol Biol. 1991, *222(3)*, S.465-77,

Radice, 1994 Radice AD, Bugaj B, Fitch DH, Emmons SW: Widespread occurrence of the Tc1 transposon family: Tc1-like transposons from teleost fish. Mol Gen Genet. 1994, *244(6)*, S.606-12,

Richardson, 2006 Richardson JM, Dawson A, O'Hagan N, Taylor P, Finnegan DJ, Walkinshaw MD: Mechanism of Mos1 transposition: insights from structural analysis. EMBO J. 2006, *25(6)*, S.1324-34,

Roberts, 1985 Roberts D, Hoopes BC, McClure WR, Kleckner N: IS10 transposition is regulated by DNA adenine methylation. Cell. 1985, *43(1)*, S.117-30,

Robertson, 1997 Robertson HM, Zumpano KL: Molecular evolution of an ancient mariner transposon, Hsmar1, in the human genome. Gene. 1997, *205(1-2)*, S.203-17,

Rogers, 1986 Rogers M, Ekaterinaki N, Nimmo E, Sherratt D: Analysis of Tn7 transposition. Mol Gen Genet. 1986, *205(3)*, S.550-6,

Savilahti, 1996 Savilahti H, Mizuuchi K: Mu transpositional recombination: donor DNA cleavage and strand transfer in trans by the Mu transposase. Cell. 1996, *85(2)*, S.271-80,

Schläppi, 1993 Schläppi M, Smith D, Fedoroff N: TnpA trans-activates methylated maize Suppressor-mutator transposable elements in transgenic tobacco. Genetics. 1993, *133(4)*, S.1009-21,

Schläppi, 1994 Schläppi M, Raina R, Fedoroff N: Epigenetic regulation of the maize Spm transposable element: novel activation of a methylated promoter by TnpA. Cell. 1994, *77(3)*, S.427-37,

Schröder, 2002 Schröder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F: HIV-1 integration in the human genome favors active genes and local hotspots. Cell. 2002, *110(4)*, S.521-9,

Sekine, 1989 Sekine Y, Ohtsubo E: Frameshifting is required for production of the transposase encoded by insertion sequence 1. Proc Natl Acad Sci U S A. 1989, *86(12)*, S.4609-13,

Signon, 1995 Signon L, Kleckner N: Negative and positive regulation of Tn10/IS10-promoted recombination by IHF: two distinguishable processes inhibit transposition off of multicopy plasmid replicons and activate chromosomal events that favor evolution of new transposons. Genes Dev. 1995, *9(9)*, S.1123-36,

Slotkin, 2009 Slotkin RK, Vaughn M, Borges F, Tanurdzić M, Becker JD, Feijó JA, Martienssen RA: Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell. 2009, *136(3)*, S.461-72,

Spanopoulou, 1996 Spanopoulou E, Zaitseva F, Wang FH, Santagata S, Baltimore D, Panayotou G: The homeodomain region of Rag-1 reveals the parallel mechanisms of bacterial and V(D)J recombination. Cell. 1996, *87(2)*, S.263-76,

Staller, 2001 Staller P, Peukert K, Kiermaier A, Seoane J, Lukas J, Karsunky H, Möröy T, Bartek J, Massagué J, Hänel F, Eilers M: Repression of p15INK4b expression by Myc through association with Miz-1. Nat Cell Biol. 2001, *3(4)*, S.392-9,

Surette, 1989 Surette MG, Lavoie BD, Chaconas G: Action at a distance in Mu DNA transposition: an enhancer-like element is the site of action of supercoiling relief activity by integration host factor (IHF). EMBO J. 1989, *8(11)*, S.3483-9,

Tang, 2005 Tang M, Cecconi C, Kim H, Bustamante C, Rio DC: Guanosine triphosphate acts as a cofactor to promote assembly of initial P-element transposase-DNA synaptic complexes. Genes Dev. 2005, *19(12)*, S.1422-5,

Tang, 2007 Tang M, Cecconi C, Bustamante C, Rio DC: Analysis of P element transposase protein-DNA interactions during the early stages of transposition. J Biol Chem. 2007, *282(39)*, S.29002-12,

Thomas, 2001 Thomas JO, Travers AA: HMG1 and 2, and related 'architectural' DNA-binding proteins. Trends Biochem Sci. 2001, *26(3)*, S.167-74,

Tosi, 2000 Tosi LR, Beverley SM: cis and trans factors affecting Mos1 mariner evolution and transposition in vitro, and its potential for functional genomics. Nucleic Acids Res. 2000, *28(3)*, S.784-90,

van Luenen, 1994 van Luenen HG, Colloms SD, Plasterk RH: The mechanism of transposition of Tc3 in C. elegans. Cell. 1994, *79(2)*, S.293-301,

Vastenhouw, 2004 Vastenhouw NL, Plasterk RH: RNAi protects the Caenorhabditis elegans germline against transposition. Trends Genet. 2004, *20(7)*, S.314-9,

Vigdal, 2002 Vigdal TJ, Kaufman CD, Izsvák Z, Voytas DF, Ivics Z: Common physical properties of DNA affecting target site selection of sleeping

beauty and other Tc1/mariner transposable elements. J Mol Biol. 2002, *323(3)*, S.441-52,

Voigt, 2008 Voigt K, Izsvák Z, Ivics Z: Targeted gene insertion for molecular medicine. J Mol Med. 2008, *86(11)*, S.1205-19,

Volff, 2006 Volff JN: Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. Bioessays. 2006, *28(9)*, S.913-22,

Vos, 1994 Vos JC , Plasterk RK: Tc1 transposase of Caenorhabditis elegans is an endonuclease with a bipartite DNA binding domain. EMBO J. 1994, *13(24)*, S.6125-32,

Vos, 1996 Vos JC, De Baere I, Plasterk RK: Transposase is the only nematode protein required for in vitro transposition of Tc1. Genes Dev. 1996, *10(6)*, S.755-61,

Walisko, 2006 Walisko O, Izsvák Z, Szabó K, Kaufman CD, Herold S, Ivics Z: Sleeping Beauty transposase modulates cell-cycle progression through interaction with Miz-1. Proc Natl Acad Sci U S A. 2006, *103(11)*, S.4062-7,

Walisko, 2008 Walisko O, Schorn A, Rolfs F, Devaraj A, Miskey C, Izsvák Z, Ivics Z: Transcriptional activities of the Sleeping Beauty transposon and shielding its genetic cargo with insulators. Mol Ther. 2008, *16(2)*, S.359-69,

Wang, 1999 Wang H, Hartswood E, Finnegan DJ: Pogo transposase contains a putative helix-turn-helix DNA binding domain that recognises a 12 bp sequence within the terminal inverted repeats. Nucleic Acids Res. 1999, *27(2)*, S.455-61,

Weinreich, 1994 Weinreich MD, Mahnke-Braam L, Reznikoff WS: A functional analysis of the Tn5 transposase. Identification of domains required for DNA binding and multimerization. J Mol Biol. 1994, *241(2)*, S.166-77,

Williams, 2008 Williams DA: Sleeping beauty vector system moves toward human trials in the United States. Mol Ther. 2008, *16(9)*, S.1515-6,

Wright, 1991 Wright WE, Binder M, Funk W: Cyclic amplification and selection of targets (CASTing) for the myogenin consensus binding site. Mol Cell Biol. 1991, *11(8)*, S.4104-10,

Wu, 2004 Wu X, Burgess SM: Integration target site selection for retroviruses and transposable elements. Cell Mol Life Sci. 2004, *61(19-20)*, S.2588-96,

Yang, 2006 Yang N, Kazazian HH Jr: L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat Struct Mol Biol. 2006, *13(9)*, S.763-71,

Yant, 2000 Yant SR, Meuse L, Chiu W, Ivics Z, Izsvak Z, Kay MA: Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. Nat Genet. 2000, *25(1)*, S.35-41,

Yin, 1988 Yin JC, Krebs MP, Reznikoff WS: Effect of dam methylation on Tn5 transposition. J Mol Biol. 1988, *199(1)*, S.35-45,

York, 1996 York D, Reznikoff WS: Purification and biochemical analyses of a monomeric form of Tn5 transposase. Nucleic Acids Res. 1996, *24(19)*, S.3790-6,

Zayed, 2003 Zayed H, Izsvák Z, Khare D, Heinemann U, Ivics Z: The DNA-bending protein HMGB1 is a cellular cofactor of Sleeping Beauty transposition. Nucleic Acids Res. 2003, *31(9)*, S.2313-22,

Zayed, 2004 Zayed H, Izsvák Z, Walisko O, Ivics Z: Development of hyperactive sleeping beauty transposon vectors by mutational analysis. Mol Ther. 2004, *9(2)*, S.292-304,

Zhang, 2001 Zhang L, Dawson A, Finnegan DJ: DNA-binding activity and subunit interaction of the mariner transposase. Nucleic Acids Res. 2001, *29(17)*, S.3566-75,

Publications and conferences

1. Pryputniewicz-Drobinska D., Kaufman C.D., Ivics Z., Izsvák Z. Regulated complex assembly protects cells from aberrant *Sleeping Beauty* transposition events. (in preparation)
2. Mátés L, Chuah .M.K.L., Belay E, Jerchow B, Manoj N, Acosta-Sanchez A, Grzela D, Schmitt A, Becker K, Matrai J, Ma L, Samara-Kuko E, Gysemans C, Prypu niewicz-Drobinska D, Miskey C, Fletcher B, VandenDriessche T, Ivics Z, Izsvák Z. Molecular Evolution of a Novel Hyperactive Sleeping Beauty Transposase Enables Robust Stable Gene Transfer in Vertebrates. Nature Genetics, 2009
3. Macierzanka M, Plotka M, Pryputniewicz-Drobinska D, Lewandowska A, Lightowlers R, Marszalek J. Maintenance and stabilization of mtDNA can be facilitated by the DNA-binding activity of Ilv5p. Biochim Biophys Acta. 2008

Presentations:

1. International Conference on Transposition and Animal Biotechnology in Berlin 2008. Titel: "Ordered complex assembly of *Sleeping Beauty* transposition"

Posters:

1. International Congress on Transposable Elements in St. Malo, France 2008
2. PhD Retreats 2006 und 2008

Erklärung / Statement

Hiermit erkläre ich, die vorliegende Arbeit selbständig angefertigt zu haben. Ich habe keine unerlaubten oder unerwähnten Hilfsmittel benutzt.

Diana Pryputniewicz-Drobinska
Berlin, September 2009